DIGITAL FORENSICS

Dr.Jeetendra Pande

Dr. Ajay Prasad

Post-Graduate Diploma in Cyber Security

Digital Forensics

(PGDCS-07)

Title	Digital Forensics
Advisors	**Mr. R. Thyagarajan**, Head, Admn. & Finance and Acting Director, CEMCA **Dr. Manas Ranjan Panigrahi**, Program Officer (Education), CEMCA **Prof. Durgesh Pant**, Director-SCS&IT, UOU
Editor	**Er. Gopesh Pande**, Network Engineer, Wipro Infotech, Mumbai
Authors	
Block I> Unit I, Unit II, Unit III & Unit IV	**Dr. Jeetendra Pande**, Assistant Professor-School of CS & IT, Uttarakhand Open University, Haldwani
Block II> Unit I, Unit II, Unit III & Unit IV	**Dr. Ajay Prasad**, Assistant Professor, CIT, University of Petroleum and Energy Studies, Dehradun
Block III> Unit I, Unit II, Unit III & Unit IV	**Dr. Ajay Prasad**, Assistant Professor, CIT, UPES, Dehradun

ISBN: 978-93-84813-94-9

Acknowledgement

The University acknowledges with thanks the expertise and financial support provided by Commonwealth Educational Media Centre for Asia (CEMCA), New Delhi, for the preparation of this study material.

Uttarakhand Open University, 2016

© Uttarakhand Open University, 2016. Digital Forensics is made available under a Creative Commons Attribution Share-Alike 4.0 Licence (international):
http://creativecommons.org/licenses/by-sa/4.0/
It is attributed to the sources marked in the References, Article Sources and Contributors section.

Published by: Uttarakhand Open University, Haldwani

Expert Panel	
S. No.	**Name**
1	**Dr. Jeetendra Pande, School of Computer Science & IT, Uttarakhand Open University, Haldwani**
2	**Prof. Ashok Panjwani, Professor, MDI, Gurgoan**
3	**Group Captain Ashok Katariya, Ministry of Defense, New Delhi**
4	**Mr. Ashutosh Bahuguna, Scientist-CERT-In, Department of Electronics & Information Technology, Government of India**
5	**Mr. Sani Abhilash, Scientist-CERT-In, Department of Electronics & Information Technology, Government of India**
6	**Wing Commander C.S. Chawla, Ministry of Defense, New Delhi**
7	**Mr. Mukesh Kumar Verma, IT Consultant, Chandigarh**
8	**Mr. Pritam Dutt Gautam, IT Consultant, New Delhi**

INDEX

1.1 LEARNING OBJECTIVES 1

1.2 INTRODUCTION 1

1.2.1 Definition of Computer Forensics 1

1.2.2 Cyber crime 1

1.2.2.1 Computer based crime 2

1.2.2.2 Computer facilitated crime 2

1.3 EVOLUTION OF COMPUTER FORENSICS 2

1.4 STAGES OF COMPUTER FORENSICS PROCESS 4

1.5 BENEFITS OF COMPUTER FORENSICS 4

1.6 USES OF COMPUTER FORENSICS 5

1.7 OBJECTIVES OF COMPUTER FORENSICS 6

1.8 ROLE OF FORENSICS INVESTIGATOR 6

1.9 FORENSICS READINESS 7

1.9.1 What is Forensics Readiness? 7

1.9.2 Goals of Forensic Readiness 7

1.9.3 Benefits of Forensic Readiness 7

1.9.4 Steps for Forensic Readiness Planning 8

1.10 SUMMARY 14

1.11 CHECK YOUR PROGRESS 15

1.12 ANSWERS TO CHECK YOUR PROGRESS 15

1.13 MODEL QUESTIONS 16

2.1 LEARNING OBJECTIVES 17

2.2 INTRODUCTION TO COMPUTER CRIME INVESTIGATION 17

2.2.1 Initial Decision-Making Process 18

2.3 ASSESS THE SITUATION 18

2.3.1 Notify Decision Makers and Acquire Authorization 18

2.3.2 Review Policies and Laws 19

2.3.3 Identify Investigation Team Members 20

2.3.4 Conduct a Thorough Assessment 21

2.3.5 Prepare for Evidence Acquisition 22

2.4 ACQUIRE THE DATA 23

2.4.1 Build Computer Investigation Toolkit 23

2.4.1.1 Preparing Your Organization for a Computer Investigation 24
2.4.2 Collect the Data 24
2.4.3 Store and Archive 27
2.5 ANALYZE THE DATA 27
2.5.1 Analyze Network Data 28
2.5.2 Analyze Host Data 28
2.5.3 Analyze Storage Media 28
2.6 REPORT THE INVESTIGATION 30
2.6.1 Gather and Organize Information 30
2.6.2 Write the Report 31
2.7 SUMMARY 32
2.8 CHECK YOUR PROGRESS 33
2.9 ANSWERS TO CHECK YOUR PROGRESS 33
2.10 MODEL QUESTIONS 34
3.1 LEARNING OBJECTIVES 35
3.2 DIGITAL EVIDENCE 35
3.2.1 Locard's Principle 35
3.2.2 Best Evidence Rule 36
3.2.3 Characteristics of Digital Evidence 36
3.2.4 Stages in Digital Evidence Investigation Process 37
3.3 FIRST RESPONDER TOOLKIT 39
3.3.1 Some Common Mistakes First Responder should avoid 40
3.4 ISSUES FACING COMPUTER FORENSICS 40
3.4.1 Technical issues 41
3.4.2 Legal issues 41
3.4.3 Administrative issues 41
3.5 TYPES OF INVESTIGATION 42
3.5.1 Criminal forensics 42
3.5.2 Intelligence gathering 42
3.5.3 Electronic discovery (eDiscovery) 42
3.5.4 Intrusion investigation 42
3.6 TECHNIQUES OF DIGITAL FORENSICS 42
3.6.1 Cross-drive analysis 42
3.6.2 Live analysis 43

3.6.2.1 Volatile data 43
3.6.3 Recovery of Deleted files 44
3.6.4 Stochastic forensics 44
3.6.5 Steganography 44
3.7 SUMMARY 44
3.8 CHECK YOUR PROGRESS 45
3.9 ANSWERS TO CHECK YOUR PROGRESS 46
3.10 MODEL QUESTIONS 47
4.1 LEARNING OBJECTIVES 48
4.2 HARD DISK DRIVE 48
4.2.1 Working 48
4.2.2 Interface 50
4.3 DETAILS OF INTERNAL STRUCTURE OF HDD 52
4.3.1 Low-Level Formatting 52
4.3.2 High-level formatting 53
4.3.3 Glossary of some important terms 54
4.3.3.1 Slack space 54
4.3.3.2 Lost Cluster 55
4.3.3.3 Bad Sector 55
4.3.3.4 Master Boot Record 55
4.4 THE BOOTING PROCESS 55
4.4.1 Linux Boot Process 58
4.4.2 Mac OS Boot Sequence 60
4.4.3 Boot Sequence in Windows 7 62
4.5 FILE SYSTEM 63
4.5.1 Some Common File systems 64
4.5.1.1 FAT 64
4.5.1.2 NTFS 64
4.5.1.3 ext2, ext3 and ext4 65
4.5.1.4 XFS 65
4.5.1.5 ZFS 65
4.5.1.6 BTRFS 65
4.5.2 Types of file systems 66
4.5.2.1 Disk file systems 66

4.5.2.2 Flash file systems 66
4.5.2.3 Tape file systems 66
4.5.2.4 Database file systems 67
4.5.2.5 Transactional file systems 68
4.5.2.6 Network file systems 69
4.5.2.7 Shared disk file systems 69
4.5.2.8 Special file systems 69
4.5.2.9 Minimal file system / audio-cassette storage 69
4.5.2.10 Flat file systems 70
4.6 SUMMARY 70
4.7 CHECK YOUR PROGRESS 71
4.8 ANSWERS TO CHECK YOUR PROGRESS 72
4.9 MODEL QUESTIONS 73
References, Article Source & Contributors 73
Further Reading 79
BLOCK II 81
1.1 LEARNING OBJECTIVES 82
1.2 INTRODUCTION 82
1.2.1 Background and need for Window forensics 82
1.2.2 Major forensic areas in windows 83
1.2.2.1 Volatile information 83
1.2.2.2 Non Volatile information 88
1.3 RECOVERING DELETED FILES AND PARTITIONS 91
1.3.1 Anatomy of a disc drive 91
1.3.2 Data organization in Windows 93
1.3.3 Retrieving deleted files 93
1.3.4 Retrieving cached files 93
1.3.5 Retrieving files in unallocated space 93
1.4 MORE ABOUT RECOVERING LOST FILES/DATA 94
1.4.1 Slack space, swap file, deleted files 94
1.4.1.1 Slack Space 94
1.4.1.2 Swap space 94
1.4.1.3 File Carving 95
1.4.1.4 Event logs 96

1.5 SUMMARY 97
1.6 CHECK YOUR PROGRESS 98
1.7 ANSWERS TO CHECK YOUR PROGRESS 98
1.8 FURTHER READINGS 99
1.9 MODEL QUESTIONS 99
References, Article Source & Contributors 99
Bibliography 99
2.1 LEARNING OBJECTIVES 100
2.2 INTRODUCTION 100
2.3 WINDOWS REGISTRY 101
2.3.1 Registry and forensics 101
2.3.1.1 System information 101
2.4 WINDOWS EVENT LOG FILE 103
2.4.1 Windows Event Log File Format 104
2.4.2 Reading from an Windows event log file 105
2.4.3 Using Microsoft log parser 106
2.4.4 Understanding Windows user account management logs 107
2.4.5 Understanding Windows file and other object Access sets 108
2.4.6 Auditing policy change 108
2.5 WINDOWS PASSWORD STORAGE 108
2.5.1 SAM 108
2.5.1.1 Removing LM hash 109
2.5.1.2 Related attacks 109
2.5.2 AD 109
2.6 APPLICATION PASSWORDS CRACKERS 110
2.6.1 Password cracking methods 110
2.6.1.1 Brute force attack 110
2.6.1.2 Dictionary attack 110
2.6.1.3 Syllable attack 111
2.6.1.4 Rule Based Attack 111
2.6.1.5 Hybrid attack and password guessing 111
2.6.1.6 Rainbow Attacks 111
2.6.1.7 System passwords 112
2.6.2 Tools for passwords cracking 112

2.6.2.1 CMOSPwd 112
2.6.2.2 ERDCommander 113
2.6.2.3 Office pwd recovery 113
2.6.2.4 Passware kit 114
2.6.2.5 PDF Password Crackers 115
2.7 SUMMARY 115
2.8 CHECK YOUR PROGRESS 116
2.9 ANSWERS TO CHECK YOUR PROGRESS 116
2.10 FURTHER READINGS 117
References, Article Source & Contributors 117
3.1 LEARNING OBJECTIVES 118
3.2 INTRODUCTION 118
3.3 NETWORK COMPONENTS AND THEIR FORENSICS IMPORTANCE 118
3.3.1 Host 119
3.3.2 Node 120
3.3.3 Router 120
3.3.4 Switch 121
3.3.5 Hub 121
3.3.6 Network interface card (NIC) 121
3.4 OSI 122
3.4.1 OSI model 122
3.4.2 TCP/IP Layers 122
3.5 FORENSICS INFORMATION FROM NETWORK 126
3.5.1 Intrusion detection/ prevention system 126
3.5.2 Wireless Access Points 127
3.6 LOG ANALYSIS 127
3.6.1 Analyzing time stamps 128
3.6.2 Analyzing data 128
3.7 FORENSICS TOOLS 128
3.7.1 Network tools used for forensics 129
3.7.2 Software tools used for network forensics 130
3.8 SUMMARY 132
3.9 CHECK YOUR PROGRESS 132
3.10 ANSWERS TO CHECK YOUR PROGRESS 133

3.11 FURTHER READINGS 134
3.12 MODEL QUESTIONS 134
References, Article Source & Contributors 134
4.1 LEARNING OBJECTIVES 136
4.2 INTRODUCTION 136
4.3WIRELESS FIDELTY (WI-FI)(802.11) 136
4.3.1 Capturing 802.11 frames 137
4.4 WIRELESS SECURITY 139
4.4.1 Attacks 140
4.5 WIRELESS ATTACKS DETECTION TECHNIQUES 142
4.5.1 Wireless Access Point Monitoring 142
4.5.2 Wireless Client/Node Monitoring 142
4.5.3 General Wireless Traffic Monitoring 143
4.6 WIRELESS INTRUSION DETECTION SYSTEMS 143
4.6.1 Snort-wireless 143
4.6.2 WIDZ 143
4.6.3 Bro 143
4.7 SUMMARY 144
4.8 CHECK YOUR PROGRESS 145
4.9 ANSWERS TO CHECK YOUR PROGRESS 145
4.10 FURTHER READING 146
References, Article Source & Contributors 146
BLOCK III 148
1.1 LEARNING OBJECTIVES 149
1.2 INTRODUCTION 149
1.2.1 Cyber-attack 149
1.2.2 Cyber Warfare and cyber terrorism 149
1.3 TYPES OF WEB ATTACKS 150
1.3.1 Spoofing 150
1.3.1.1 Email spoofing 150
1.3.1.2 Website spoofing 151
1.3.2 Repudiation 151
1.3.3 Privacy attack 152
1.3.4 Denial of Service 153

1.3.5 Privilege escalation 153
1.3.6 SQL Injection Attacks 154
1.4 WEB ATTACK FORENSICS 154
1.4.1 Web services forensics 155
1.4.2 Web Application Forensics 155
1.4.3 Preliminary Analysis 156
1.4.3.1 Application Forensics Readiness 156
1.4.3.2 Methodology 156
1.4.4 Website traffic analysis 157
1.5 WEB APPLICATION FORENSICS TOOLS 158
1.5.1 Logparser 158
1.5.2 EventLog Analyzer 158
1.5.3 Web log analyzer 159
1.5.4 Open Web Analytics[8] 159
1.5.5 Webalizer 159
1.6 SUMMARY 159
1.7 CHECK YOUR PROGRESS 159
1.8 ANSWERS TO CHECK YOUR PROGRESS 160
1.9 FURTHER READINGS 160
1.10 MODEL QUESTIONS 161
References, Article Source & Contributors 161
2.1 LEARNING OBJECTIVES 162
2.2 INTRODUCTION 162
2.2.1 Email Structure 162
2.2.2 Types of email services 162
2.2.2.1 Web-based email 162
2.2.2.2 POP3 email services 162
2.2.2.3 The Internet Message Access Protocol (IMAP) 163
2.2.2.4 MAPI email servers 163
2.2.3 Simple Mail Transfer Protocol 163
2.3 EMAIL ATTACKS AND CRIMES 164
2.3.1 Flaming 165
2.3.2 Email spoofing 165
2.3.3 Email bombing 165

2.3.4 Email hacking 165
2.3.5 Spams 165
2.3.6 Phishing 165
2.3.7 Email fraud 165
2.3.8 Phishing emails 166
2.4 PRIVACY IN EMAILS 166
2.4.1 Email privacy 166
2.4.2 Email tracking 167
2.5 EMAIL FORENSICS 167
2.5.1 Forensically important email parts 167
2.5.2 Email forensics investigation 169
2.5.3 Analyzing an email 170
2.5.4 Instant Messages 174
2.6 EMAIL FORENSIC TOOLS 175
2.6.1 eMailTrackerPro 175
2.6.2 Online EMailTracer 176
2.7 SUMMARY 176
2.8 CHECK YOUR PROGRESS 177
2.9 ANSWERS TO CHECK YOUR PROGRESS 177
2.10 MODEL QUESTIONS 178
2.11 FURTHER READINGS 178
References, Article Source & Contributors 178
3.1 LEARNING OBJECTIVES 179
3.2 INTRODUCTION 179
3.3 CHALLENGES IN MOBILE FORENSICS 179
3.4 MOBILE COMMUNICATION 180
3.4.1 802.11 or WiFi 180
3.4.2 Bluetooth 180
3.4.3 Infrared (IrDA) 181
3.5 EVIDENCES IN A MOBILE DEVICE 181
3.5.1 Service provider logs 181
3.5.2 Subscriber identification module 182
3.5.3 Mobile Logs 182
3.5.4 Phone books/contact lists 182

3.5.5 Text messages 182
3.5.6 Application files 182
3.6 MOBILE FORENSIC PROCESS 182
3.6.1 Seizure 183
3.6.2 Acquisition 183
3.6.3 Examination and Analysis 185
3.7 FORENSIC ACQUISITION TOOLS 185
3.7.1 Hardware acquisition tools 185
3.7.2 Software acquisition tools 186
3.8 SUMMARY 188
3.9 CHECK YOUR PROGRESS 188
3.10 ANSWERS TO CHECK YOUR PROGRESS 189
3.11 FURTHER READINGS 189
3.12 MODEL QUESTIONS 190
References, Article Source & Contributors 190
4.1 LEARNING OBJECTIVES 191
4.2 INTRODUCTION 191
4.3 REPORT PREPARATION 192
4.3.1 Gathering the data 192
4.3.2 Analyzing the results 192
4.3.3 Outlining and organizing the report 193
4.3.4 Writing and Revising a Rough Draft 194
4.4 Expert Witness 195
4.4.1 Finding an expert 196
4.4.1.1 Testifying v. Consulting 196
4.4.2 What Can (and Can't) an Expert Do? 196
4.4.3 Why Use an Expert 197
4.5 LEGAL ASPECTS OF COMPUTING 197
4.5.1 Jurisdiction 198
4.5.2 Net neutrality 199
4.5.3 Open Internet 199
4.5.4 Indian Information Technology Act(IT Act) 2000[6] 200
4.5.4.1 Against Individual 201
4.5.4.2 Individual Property 201

4.5.4.3 Against Organisation 201
4.5.4.4 Against Society at Large 201
4.5.5 Amendments– Indian IT Act (2008) 201
4.6 SUMMARY 206
4.7 CHECK YOUR PROGRESS 206
4.8 ANSWERS TO CHECK YOUR PROGRESS 207
4.9 FURTHER READINGS 207
4.10 MODEL QUESTIONS 207
References, Article Source & Contributors 207

BLOCK I

UNIT I: INTRODUCTION TO DIGITAL FORENSIC

1.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Know the history and evolution of digital forensics
- Describe various types of cyber crime
- Understand benefits of computer forensics
- Know about forensics readiness
- Implement forensics readiness plan

1.2 INTRODUCTION

Digital forensics[1], the art of recovering and analysing the contents found on digital devices such as desktops, notebooks/netbooks, tablets, smartphones, etc., was little-known a few years ago. However, with the growing incidence of cyber crime, and the increased adoption of digital devices, this branch of forensics has gained significant importance in the recent past, augmenting what was conventionally limited to the recovery and analysis of biological and chemical evidence during criminal investigations.

1.2.1 Definition of Computer Forensics

Computer forensics[2] is the practice of collecting, analysing and reporting on digital data in a way that is legally admissible. It can be used in the detection and prevention of crime and in any dispute where evidence is stored digitally. It is the use of specialized techniques for recovery, authentication and analysis of electronic data when a case involves issues relating to reconstruction of computer usage, examination of residual data, and authentication of data by technical analysis or explanation of technical features of data and computer usage[3]. Computer forensics requires specialized expertise that goes beyond normal data collection and preservation techniques available to end-users or system support personnel. Similar to all forms of forensic science, computer forensics is comprised of the application of the law to computer science. Computer forensics deals with the preservation, identification, extraction, and documentation of computer evidence. Like many other forensic sciences, computer forensics involves the use of sophisticated technological tools and procedures that must be followed to guarantee the accuracy of the preservation of evidence and the accuracy of results concerning computer evidence processing. The use of specialized techniques for recovery, authentication, and analysis of computer data, typically of data which may have been deleted or destroyed.

1.2.2 Cyber crime

Computer crime[4], or **cybercrime**, is any crime that involves a computer and a network. The computer may have been used in the commission of a crime, or it may be the target. Dr.

[1] http://opensourceforu.efytimes.com/2011/03/digital-forensic-analysis-using-backtrack-part-1/
[2] https://forensiccontrol.com/resources/beginners-guide-computer-forensics/
[3] http://www.edrm.net/resources/glossaries/glossary/c/computer-forensics
[4] https://en.wikipedia.org/wiki/Computer_crime

Debarati Halder and Dr. K. Jaishankar define Cybercrimes as: "Offences that are committed against individuals or groups of individuals with a criminal motive to intentionally harm the reputation of the victim or cause physical or mental harm, or loss, to the victim directly or indirectly, using modern telecommunication networks such as Internet (Chat rooms, emails, notice boards and groups) and mobile phones (SMS/MMS)". Such crimes may threaten a nation's security and financial health. Issues surrounding these types of crimes have become high-profile, particularly those surrounding hacking, copyright infringement, child pornography, and child grooming. There are also problems of privacy when confidential information is intercepted or disclosed, lawfully or otherwise.

Internationally, both governmental and non-state actors engage in cybercrimes, including espionage, financial theft, and other cross-border crimes. Activity crossing international borders and involving the interests of at least one nation state is sometimes referred to as cyberwarfare.

Digital forensics is traditionally associated with criminal investigations and, as you would expect, most types of investigation centre on some form of computer crime. This sort of crime can take two forms; computer based crime and computer facilitated crime[5].

1.2.2.1 Computer based crime

This is criminal activity that is conducted purely on computers, for example cyber-bullying or spam. As well as crimes newly defined by the computing age it also includes traditional crime conducted purely on computers (for example, child pornography).

1.2.2.2 Computer facilitated crime

Crime conducted in the "real world" but facilitated by the use of computers. A classic example of this sort of crime is fraud: computers are commonly used to communicate with other fraudsters, to record/plan activities or to create fraudulent documents.

Not all digital forensics investigations focus on criminal behaviour; sometimes the techniques are used in corporate (or private) settings to recover lost information or to rebuild the activities of employees.

1.3 EVOLUTION OF COMPUTER FORENSICS

It is difficult to pinpoint the first "computer forensic" examination or the beginning of the field for that matter[6]. But most experts agree that the field of computer forensics began to evolve more than 30 years ago. The field began in the United States, in large part, when law enforcement and military investigators started seeing criminals get technical. Government personnel charged with protecting important, confidential, and certainly secret information conducted forensic examinations in response to potential security breaches to not only investigate the particular breach, but to learn how to prevent future potential breaches. Ultimately, the fields of information security, which focuses on protecting information and assets, and computer forensics, which focuses on the response to hi-tech offenses, started to intertwine.

[5] https://en.wikibooks.org/wiki/Introduction_to_Digital_Forensics/Types
[6] http://colbycriminaljustice.wikidot.com/cyberforensics

Over the next decades, and up to today, the field is evolving. Both Government and private organizations and corporations have followed suit – employing internal information security and computer forensic professionals or contracting such professionals or firms on an as-needed basis. Significantly, the private legal industry has more recently seen the need for computer forensic examinations in civil legal disputes, causing an explosion in the e-discovery field.

The history of forensic science dates back thousands of years. Fingerprinting was one of its first applications. The ancient Chinese used fingerprints to identify business documents. In 1892, a eugenicist named Sir Francis Galton established the first system for classifying fingerprints. Sir Edward Henry, commissioner of the Metropolitan Police of London, developed his own system in 1896 based on the direction, flow, pattern and other characteristics in fingerprints. The Henry Classification System became the standard for criminal fingerprinting techniques worldwide.

In 1835, Scotland Yard's Henry Goddard became the first person to use physical analysis to connect a bullet to the murder weapon. Bullet examination became more precise in the 1920s, when American physician Calvin Goddard created the comparison microscope to help determine which bullets came from which shell casings. And in the 1970s, a team of scientists at the Aerospace Corporation in California developed a method for detecting gunshot residue using scanning electron microscopes.

In 1836, a Scottish chemist named James Marsh developed a chemical test to detect arsenic, which was used during a murder trial. Nearly a century later, in 1930, scientist Karl Landsteiner won the Nobel Prize for classifying human blood into its various groups. His work paved the way for the future use of blood in criminal investigations. Other tests were developed in the mid-1900s to analyze saliva, semen and other body fluids as well as to make blood tests more precise.

In 1984, FBI Magnetic Media program, which was later renamed to Computer Analysis and Response Team (CART), was created and it is believed to be the beginning of computer forensic.

In 1988, the International Association of Computer Investigative Specialists(IACIS), an international non-profit corporation composed of volunteer computer forensic professionals dedicated to training and certifying practitioners in the field of forensic computer science was formed.

It was followed by formation of International Organization on Computer Evidence(IOCE)[7] in 1995, which aims to brings together organizations actively engaged in the field of digital and multimedia evidence to foster communication and cooperation as well as to ensure quality and consistency within the forensic community.

With the rise in cyber crime, the G8 nations realised the importance of computer forensic, and in 1997 declared that “Law enforcement personnel must be trained and equipped to address

[7] https://en.wikipedia.org/wiki/Scientific_Working_Group_on_Digital_Evidence

high-tech crimes". In 1998, G8 appointed IICE to create international principles, guidelines and procedures relating to digital evidence. In the same year INTERPOL Forensic Science Symposium was held. The First FBI Regional Computer Forensic Laboratory established in 2000 at San Diego.
The timeline of computer forensics could be summarised as:

Table 1: Computer Forensics Timeline

Year	Event
1835	Scotland Yard's Henry Goddard became the first person to use physical analysis to connect a bullet to the murder weapon.
1836	James Marsh developed a chemical test to detect arsenic, which was used during a murder trial.
1892	Sir Francis Galton established the first system for classifying fingerprints.
1896	Sir Edward Henry, based on the direction, flow, pattern and other characteristics in fingerprints.
1920	American physician Calvin Goddard created the comparison microscope to help determine which bullets came from which shell casings.
1930	Karl Landsteiner won the Nobel Prize for classifying human blood into its various groups.
1970	Aerospace Corporation in California developed a method for detecting gunshot residue using scanning electron microscopes.
1984	FBI Magnetic Media program, which was later renamed to Computer Analysis and Response Team (CART), was created and it is believed to be the beginning of computer forensic.
1988	International Association of Computer Investigative Specialists(IACIS) was formed.
1995	International Organization on Computer Evidence (IOCE) was formed.
1997	G8 nations declared that "Law enforcement personnel must be trained and equipped to address high-tech crimes".
1998	• G8 appointed IICE to create international principles, guidelines and procedures relating to digital evidence. • 1st INTERPOL Forensic Science Symposium was held.
2000	First FBI Regional Computer Forensic Laboratory established.

1.4 STAGES OF COMPUTER FORENSICS PROCESS

The overall computer forensics process is sometimes viewed as comprising four stages[8]:

- Acquire: Identifying and Preserving
- Analyze: Technical Analysis
- Evaluate: What the Lawyers Do
- Present: Present digital evidence in a manner that is legally acceptable in any legal proceedings.

1.5 BENEFITS OF COMPUTER FORENSICS

With the ever increasing rate of cyber crimes, from phishing to hacking and stealing of personal information not only confined to a particular country but the globally at large, there is a need for forensic experts to be available in public and private organizations[9]. To be able to handle this, it's vital for network administrator and security staff of networked organizations to have this course in practice making sure that they have the laws pertaining to

[8] http://computer-forensics.safemode.org/index.php?page=Four_Step_Process

[9] http://computer-forensics.wikidot.com/

this on their finger tips. This would ensure that should need for the service avail itself, then they would come in and rescue the situation.

The survival and integrity of any given network infrastructure of any company or organization strongly depends on the application of computer forensics. They should be taken as the main element of computer and network security. It would be a great benefit for a company if it has knowledge of all the technical and legal aspects of this field. Should the company's network be under attack and the intruder caught in the act, then an understanding about computer forensics will be of help in provision of evidence and prosecution of the case in the court of law.

New laws aimed at the protection of customer's data are continuously being developed. Should they lose data, then naturally the liability goes to the company. Such cases, if they occur will automatically result in the company or organization being brought to the court of law for failure to protect personal data, this can turn out to be very expensive. But through the application of forensic science, huge chunks of money can be saved by the firms concerned. A lot of money is lately being spent on network and computer security. Software for vulnerability assessment and intrusion detection has passed the billion dollar mark, this is according to experts. It simply means that there is a necessity in investment in either employing an expert in computer forensic in the firms, or having part of their staff trained into this venture so as to help in detection of such cases should they arise.

1.6 USES OF COMPUTER FORENSICS

There are few areas of crime or dispute where computer forensics cannot be applied[2]. Law enforcement agencies have been among the earliest and heaviest users of computer forensics and consequently have often been at the forefront of developments in the field.

Computers may constitute a 'scene of a crime', for example with hacking or denial of service attacks or they may hold evidence in the form of emails, internet history, documents or other files relevant to crimes such as murder, kidnap, fraud and drug trafficking.

It is not just the content of emails, documents and other files which may be of interest to investigators but also the 'metadata' associated with those files. A computer forensic examination may reveal when a document first appeared on a computer, when it was last edited, when it was last saved or printed and which user carried out these actions.

More recently, commercial organisations have used computer forensics to their benefit in a variety of cases such as:

- Intellectual Property theft
- Industrial espionage
- Employment disputes
- Fraud investigations
- Forgeries
- Bankruptcy investigations
- Inappropriate email and internet use in the work place
- Regulatory compliance

1.7 OBJECTIVES OF COMPUTER FORENSICS

We all will agree to the fact that we are depending more on more on Information & Communication Technology(ICT) tools and internet for digital services to an extent that today we talk online using chat application, we depend on email to communicate with relatives and office, we stay in touch with our friends and update status using social engineering platforms like facebook, etc., we work online by staying connected to our office/ clinet using internet, we shop online, we teach online, we learn online, we submit our bill online today. Our dependency on Computer and Internet have increased so much that we are "online" most of the time. Therefore, there is an increased need of protecting our information from being misused by following Information security guidelines. However, if the security of our computer is compromised, computer forensics comes handy for post- incident investigation.

The objectives of Computer forensics are to provide guidelines for:

- Following the first responder procedure and access the victim's computer after incident.
- Designing procedures at a suspected crime scene to ensure that the digital evidence obtained is not corrupted.
- Data acquisition and duplication.
- Recovering deleted files and deleted partitions from digital media to extract the evidence and validate them.
- Provide guidelines for analyzing digital media to preserve evidence, analysing logs and deriving conclusions, investigate network traffics and logs to correlate events, investigate wireless and web attacks, tracking emails and investigate email crimes.
- Producing computer forensic report which provides complete report on computer forensic investigation process.
- Preserving the evidence by following the chain of custody.
- Employing the rigorous procedures necessary to have forensic results stand up to scrutiny in a court of law.
- Presenting digital forensics results in a court of law as an expert witness.

1.8 ROLE OF FORENSICS INVESTIGATOR

Following are some of the important duties of a forensic investigator:

- Confirms or dispels whether a resource/network is compromised.
- Determine extent of damage due to intrusion.
- Answer the questions: Who, What, When, Where, How and Why.
- Gathering data in a forensically sound manner.
- Handle and analyze evidence.
- Prepare the report.
- Present admissible evidence in court.

1.9 FORENSICS READINESS

There are several reasons for this field's growth; the most significant being that computers are everywhere[10]. You'd be hard pressed to find a household today without at least one computer. And it is not just computers that computer forensic examiners get involved with. Computer forensic examiners analyze all types of technical devices. Look around you while you walk down the street – people are on their cell phones, using iPods, PDAs, and text messaging. Computer forensic examiners analyze all of these electronic devices! Cyber forensics is a rapidly changing field. There are new technologies coming out daily that are becoming smaller, but storing more and more data. This leads to why cyber forensics is import. In computer related crimes, such identity fraud, it is becoming easier to hide data. With the proper analysis of digital evidence, better security can be made to protect computer users, but also catch those who are committing the crimes. Organizations have now realised the importance of being prepared to combat cyber criminals with their forensic readiness plan ready.

1.9.1 What is Forensics Readiness?

Forensic readiness is the ability of an organisation to maximise its potential to use digital evidence whilst minimising the costs of an investigation[11]. In a business context there is the opportunity to actively collect potential evidence in the form of logfiles, emails, back-up disks, portable computers, network traffic records, and telephone records, amongst others. This evidence may be collected in advance of a crime or dispute, and may be used to the benefit of the collecting organisation if it becomes involved in a formal dispute or legal process.

1.9.2 Goals of Forensic Readiness

Some of the important goals of forensics readiness are:

- to gather admissible evidence legally and without interfering with business processes;
- to gather evidence targeting the potential crimes and disputes that may adversely impact an organisation;
- to allow an investigation to proceed at a cost in proportion to the incident;
- to minimise interruption to the business from any investigation; and
- to ensure that evidence makes a positive impact on the outcome of any legal action.

1.9.3 Benefits of Forensic Readiness

Forensic readiness can offer an organisation the following benefits:

- evidence can be gathered to act in an organisation's defence if subject to a lawsuit;
- comprehensive evidence gathering can be used as a deterrent to the insider threat (throwing away potential evidence is simply helping to cover the tracks of a cyber-criminal);
- in the event of a major incident, an efficient and rapid investigation can be conducted and actions taken with minimal disruption to the business;

[10] http://colbycriminaljustice.wikidot.com/cyberforensics
[11] http://www.cpni.gov.uk/Documents/Publications/2005/2005008-TN1005_Forensic_readiness_planning.pdf

- a systematic approach to evidence storage can significantly reduce the costs and time of an internal investigation;
- a structured approach to evidence storage can reduce the costs of any court-ordered disclosure or regulatory or legal need to disclose data (e.g. in response to a request under data protection legislation);
- forensic readiness can extend the scope of information security to the wider threat from cyber crime, such as intellectual property protection, fraud, extortion etc;
- it demonstrates due diligence and good corporate governance of the company's information assets;
- it can demonstrate that regulatory requirements have been met;
- it can improve and facilitate the interface to law enforcement if involved;
- it can improve the prospects for a successful legal action;
- it can provide evidence to resolve a commercial dispute; and
- it can support employee sanctions based on digital evidence (for example to prove violation of an acceptable use policy)

1.9.4 Steps for Forensic Readiness Planning

The following ten steps describe the key activities in forensic readiness planning:

1. Define the business scenarios that require digital evidence;
2. Identify available sources and different types of potential evidence;
3. Determine the evidence collection requirement;
4. Establish a capability for securely gathering legally admissible evidence to meet the requirement;
5. Establish a policy for secure storage and handling of potential evidence;
6. Ensure monitoring is targeted to detect and deter major incidents;
7. Specify circumstances when escalation to a full formal investigation (which may use the digital evidence) should be launched;
8. Train staff in incident awareness, so that all those involved understand their role in the digital evidence process and the legal sensitivities of evidence;
9. Document an evidence-based case describing the incident and its impact; and
10. Ensure legal review to facilitate action in response to the incident.

The remainder of this section gives a brief description of each of the ten steps.

1. **Define the business scenarios that require digital evidence:** The first step in forensic readiness is to define the purpose of an evidence collection capability. The rationale is to look at the risk and potential impact on the business from the various types of crimes and disputes. What is the threat to the business and what parts are vulnerable? This is, in effect, a risk assessment, and is performed at the business level. The aim is to understand the business scenarios where digital evidence may be required and may benefit the organisation the event that it is required. In general the areas where digital evidence can be applied include:
   - reducing the impact from computer-related crime;
   - dealing effectively with court orders to release data;
   - demonstrating compliance with regulatory or legal constraints;

- producing evidence to support company disciplinary issues;
- supporting contractual and commercial agreements; and
- proving the impact of a crime or dispute.

In assessing these scenarios, this step provides an indication of the likely benefits of being able to use digital evidence. If the identified risks, and the potential benefits of forensic readiness, suggest a good return on investment is achievable, then an organisation needs to consider what evidence to gather for the various risk scenarios.

2. **Identify available sources and different types of potential evidence:** The second step in forensic readiness is for an organisation to know what sources of potential evidence are present on, or could be generated by, their systems and to determine what currently happens to the potential evidence data. Computer logs can originate from many sources. The purpose of this step is to scope what evidence may be available from across the range of systems and applications in use. Some basic questions need to be asked about possible evidence sources to include.
   - Where is data generated?
   - What format is it in?
   - How long is it stored for?
   - How is it currently controlled, secured and managed?
   - Who has access to the data?
   - How much is produced?
   - Is it archived? If so where and for how long?
   - How much is reviewed?
   - What additional evidence sources could be enabled?
   - Who is responsible for this data?
   - Who is the formal owner of the data?
   - How could it be made available to an investigation?
   - What business processes does it relate to?
   - Does it contain personal information?

   Email is an obvious example of a potential rich source of evidence that needs careful consideration in terms of storage, archiving & auditing and retrieval. But this is not the only means of communication used over the internet, there is also instant messaging, web-based email that bypasses corporate email servers, chat-rooms and newsgroups, even voice over the internet. Each of these may need preserving and archiving. The range of possible evidence sources includes:
   - equipment such as routers, firewalls, servers, clients, portables, embedded devices etc;
   - application software such as accounting packages etc for evidence of fraud, erp packages for employee records and activities (e.g. in case of identity theft), system and management files etc;
   - monitoring software such as intrusion detection software, packet sniffers, keyboard loggers, content checkers, etc;
   - general logs such as access logs, printer logs, web traffic, internal network logs, internet traffic, database transactions, commercial transactions etc;

- other sources such as: cctv, door access records, phone logs, pabx data etc; and
- back-ups and archives.

3. **Determine the Evidence Collection Requirement:** It is now possible to decide which of the possible evidence sources identified in step 2 can help deal with the crimes and disputes identified in step 1 and whether further ways to gather evidence are required. This is the evidence collection requirement. The purpose of this step is to produce an evidence requirement statement so that those responsible for managing the business risk can communicate with those running and monitoring information systems through an agreed requirement for evidence. One of the key benefits of this step is the bringing together of IT with the needs of corporate security. IT audit logs have been traditionally configured by systems administrators independently of corporate policy and where such a policy exists there is often a significant gap between organisational security objectives and the 'bottom-up' auditing actually implemented. The evidence collection requirement is moderated by a cost benefit analysis of how much the required evidence will cost to collect and what benefit it provides (see above). The critical question for successful forensic readiness is what can be performed cost effectively. By considering these issues in advance and choosing storage options, auditing tools, investigation tools, and appropriate procedures it is possible for an organisation to reduce the costs of future forensic investigations.
4. **Establish a capability for securely gathering legally admissible evidence to meet the requirement:** At this point the organisation knows the totality of evidence available and has decided which of it can be collected to address the company risks and within a planned budget. With the evidence requirement understood, the next step is to ensure that it is collected from the relevant sources and that it is preserved as an authentic record. At this stage legal advice is required to ensure that the evidence can be gathered legally and the evidence requirement can be met in the manner planned. For example, does it involve monitoring personal emails, the use of personal data, or 'fishing trips1 ' on employee activities? In some countries, some or all of these activities may be illegal. Relevant laws, in the areas of data protection, privacy and human rights, will inevitably constrain what can actually be gathered. Some of the guidelines are:
   - monitoring should be targeted at specific problems.
   - it should only be gathered for defined purposes and nothing more; and
   - staff should be told what monitoring is happening except in exceptional circumstances.

   Physical security of data such as back-up files or on central log servers is important from the data protection point of view, and also for secure evidence storage. As well as preventative measures such as secure rooms and swipe card access it is also prudent to have records of who has access to the general location and who has access to the actual machines containing evidence. Any evidence or paperwork associated with a specific investigation should be given added security by, for example, storing in a safe. Additional security of logs can also be achieved through the use of WORM storage media.

5. **Establish a policy for secure storage and handling of potential evidence:** The objective of this step is to secure the evidence for the longer term once it has been collected and to facilitate its retrieval if required. It concerns the long-term or off-line storage of information that might be required for evidence at a later date. A policy for secure storage and handling of potential evidence comprises security measures to ensure the authenticity of the data and also procedures to demonstrate that the evidence integrity is preserved whenever it is used, moved or combined with new evidence. In the parlance of investigators this is known as continuity of evidence (in the UK) and chain of custody (in the US). The continuity of evidence also includes records of who held, and who had access to, the evidence (for example from swipe control door logs). A significant contribution to the legal collection of evidence is given by the code of practice on the legal admissibility and weight of information stored electronically, published by the British Standards Institution. This document originated from a perceived need for evidence collection in the paperless office. The problem it addressed is if all paper documents are scanned, can the paper sources be thrown away without loss of evidential usability? The current edition broadens the scope to all information management systems, Ad hoc opportunistic searches, without justification, for potentially incriminating activities or communication such as those where information is transmitted over networks such as email systems for example. It points out that methods of storage, hardware reliability, operation and access control, and even the programs and source code, may be investigated in order to determine admissibility. A closely related international standard is being developed as ISO 15801. The required output of this step is a secure evidence policy. It should document the security measures, the legal advice and the procedural measures used to ensure the evidence requirement is met. Upon this document rests the likely admissibility and weight of any evidence gathered.
6. **Ensure monitoring and auditing is targeted to detect and deter major incidents:** In addition to gathering evidence for later use in court, evidence sources can be monitored to detect threatened incidents in a timely manner. This is directly analogous to Intrusion Detection Systems (IDS), extended beyond network attack to a wide range of behaviours that may have implications for the organisation. It is all very well collecting the evidence. This step is about making sure it can be used in the process of detection. By monitoring sources of evidence we can look for the triggers that mean something suspicious may be happening. The critical question in this step is when should an organisation be suspicious? A suspicious event has to be related to business risk and not couched in technical terms. Thus the onus is on managers to explain to those monitoring the data what they want to prevent and thus the sort of behaviour that IDS might be used to detect for example. This should be captured in a 'suspicion' policy that helps the various monitoring and auditing staff understand what triggers should provoke suspicion, who to report the suspicion to, whether heightened monitoring is required, and whether any additional security measures should be taken as a precaution. Each type of monitoring should produce a proportion of false positives. The sensitivity of triggers can be varied as long as the overall false positive rate does not become so high that suspicious events cannot be properly reviewed.

Varying triggers also guards against the risk from someone who knows what the threshold on a particular event is and makes sure any events or transactions he wishes to hide are beneath it.

7. **Specify circumstances when escalation to a full formal investigation (which may use digital evidence) is required:** Some suspicious events can be system generated, such as by the rule-base of an IDS, or the keywords of a content checker, and some will be triggered by human watchfulness. Each suspicious event found in step 6 needs to be reviewed. Either an event will require escalation if it is clearly serious enough, or it will require enhanced monitoring or other precautionary measures, or it is a false positive. The purpose of this step is to decide how to react to the suspicious event. The decision as to whether to escalate the situation to management will depend on any indications that a major business impact is likely or that a full investigation may be required where digital evidence may be needed. The decision criteria should be captured in an escalation policy that makes it clear when a suspicious event becomes a confirmed incident. At this point an investigation should be launched and policy should indicate who the points of contact are (potentially available on a 24x7 basis) and who else needs to be involved. As with steps 3 and 6, the network and IT security managers and the non-IT managers need to understand each other's position. What level of certainty or level of risk is appropriate for an escalation? What strength of case is required to proceed? A preliminary business impact assessment should be made based on whether any of the following are present:
   - evidence of a reportable crime
   - evidence of internal fraud, theft, other loss
   - estimate of possible damages (a threshold may induce an escalation trigger)
   - potential for embarrassment, reputation loss
   - any immediate impact on customers, partners or profitability
   - recovery plans have been enacted or are required; and
   - the incident is reportable under a compliance regime.
8. **Train staff, so that all those involved understand their role in the digital evidence process and the legal sensitivities of evidence:** A wide range of staff may become involved in a computer security incident. The aim of this step is to ensure that appropriate training is developed to prepare staff for the various roles they may play before, during and after an incident. It is also necessary to ensure that staff is competent to perform any roles related to the handling and preservation of evidence. There will be some issues relevant to all staff if they become involved in an incident. The following groups will require more specialised awareness training for example:
   - the investigating team;
   - corporate HR department;
   - corporate PR department (to manage any public information about the incident);
   - 'owners' of business processes or data;
   - line management, profit centre managers;
   - corporate security;
   - system administrators;
   - IT management;

- legal advisers; and
- senior management (potentially up to board level).

At all times those involved should act according to 'need to know' principles. They should be particularly aware whether any staff, such as 'whistle blowers' and investigators, need to be protected from possible retaliation by keeping their names and their involvement confidential. Training may also be required to understand the relationships and necessary communications with external organisations that may become involved.

9. **Present an evidence-based case describing the incident and its impact:** The aim of an investigation is not just to find a culprit or repair any damage. An investigation has to provide answers to questions and demonstrate why those answers are credible. The questions go along the lines of who, what, why, when, where and how. Credibility is provided by evidence and a logical argument. The purpose of this step is to produce a policy that describes how an evidence-based case should be assembled. A case file may be required for a number of reasons:
   - to provide a basis for interaction with legal advisers and law enforcement;
   - to support a report to a regulatory body;
   - to support an insurance claim;
   - to justify disciplinary action;
   - to provide feedback on how such an incident can be avoided in future;
   - to provide a record in case of a similar event in the future (supports the corporate memory so that even if there are changes in personnel it will still be possible to understand what has happened); and
   - to provide further evidence if required in the future, for example if no action is deemed necessary at this point but further developments occur.
10. **Ensure legal review to facilitate action in response to the incident:** At certain points during the collating of the cyber-crime case file it will be necessary to review the case from a legal standpoint and get legal advice on any follow-up actions. Legal advisers should be able to advise on the strength of the case and suggest whether additional measures should be taken; for example, if the evidence is weak is it necessary to catch an internal suspect red handed by monitoring their activity and seizing their PC? Any progression to a formal action will need to be justified, cost-effective and assessed as likely to end in the company's favour. Although the actual decision of how to proceed will clearly be post-incident, considerable legal preparation is required in readiness. Legal advisors should be trained and experienced in the appropriate cyberlaws and evidence admissibility issues. They need to be prepared to act on an incident, pursuant to the digital evidence that has been gathered and the case presented in step 9. Legal advice should also recognise that the legal issues may span legal jurisdictions e.g. states in the US, member states in the EU. Advice from legal advisers will include:
   - any liabilities from the incident and how they can be managed;
   - finding and prosecuting/punishing (internal versus external culprits);
   - legal and regulatory constraints on what action can be taken;
   - reputation protection and PR issues;

- when/if to advise partners, customers and investors;
- how to deal with employees;
- resolving commercial disputes; and
- any additional measures required.

1.10 SUMMARY

1. Computer forensics is the practice of collecting, analysing and reporting on digital data in a way that is legally admissible.
2. Computer forensics requires specialized expertise that goes beyond normal data collection and preservation techniques available to end-users or system support personnel.
3. Computer crime, or cybercrime, is any crime that involves a computer and a network.
4. Activity crossing international borders and involving the interests of at least one nation state is sometimes referred to as cyberwarfare.
5. The ancient Chinese used fingerprints to identify business documents.
6. Sir Francis Galton established the first system for classifying fingerprints.
7. International Association of Computer Investigative Specialists(IACIS) is an international non-profit corporation composed of volunteer computer forensic professionals dedicated to training and certifying practitioners in the field of forensic computer science.
8. The First FBI Regional Computer Forensic Laboratory established in 2000 at San Diego.
9. The survival and integrity of any given network infrastructure of any company or organization strongly depends on the application of computer forensics.
10. Forensic readiness is the ability of an organisation to maximise its potential to use digital evidence whilst minimising the costs of an investigation.
11. Monitoring should be targeted at specific problems.
12. Physical security of data such as back-up files or on central log servers is important from the data protection point of view, and also for secure evidence storage.
13. A policy for secure storage and handling of potential evidence comprises security measures to ensure the authenticity of the data and also procedures to demonstrate that the evidence integrity is preserved whenever it is used, moved or combined with new evidence.
14. In addition to gathering evidence for later use in court, evidence sources can be monitored to detect threatened incidents in a timely manner.
15. Some suspicious events can be system generated, such as by the rule-base of an IDS, or the keywords of a content checker, and some will be triggered by human watchfulness.
16. The decision as to whether to escalate the situation to management will depend on any indications that a major business impact is likely or that a full investigation may be required where digital evidence may be needed.
17. It is necessary to ensure that staff is competent to perform any roles related to the handling and preservation of evidence.

18. The aim of an investigation is not just to find a culprit or repair any damage. An investigation has to provide answers to questions and demonstrate why those answers are credible.
19. At certain points during the collating of the cyber-crime case file it will be necessary to review the case from a legal standpoint and get legal advice on any follow-up actions.

1.11 CHECK YOUR PROGRESS

1. Fill in the blanks

i. _____ was one of the first applications of forensics.
ii. FBI Magnetic Media program was later renamed to _____.
iii. _____is provided by evidence and a logical argument.
iv. At all times those involved should act according to _____principles.
v. IACIS stands for _____.
vi. The first step in forensic readiness is to define the _____of an evidence collection capability.
vii. It is not just the content of emails, documents and other files which may be of interest to investigators but also the _____associated with those files.
viii. IDS stands for _____.
ix. The decision criteria should be captured in an _____ policy that makes it clear when a suspicious event becomes a confirmed incident.
x. IOCE stands for International_____.

2. State true or false

i. Cybercrime, is any crime that involves a computer and a network.
ii. Computer based crime is criminal activity that is conducted purely on computers, for example cyber-bullying or spam.
iii. The goal of forensic readiness is to gather admissible evidence legally and without interfering with business processes.
iv. FBI Magnetic Media program started in 1994.
v. IOCE aims to bring together organizations actively engaged in the field of digital and multimedia evidence to foster communication and cooperation as well as to ensure quality and consistency within the forensic community.
vi. Logs can originate from only one source in a computer.
vii. The range of possible evidence sources includes equipment such as routers, firewalls, servers, clients, portables, embedded devices etc.
viii. Email is an obvious example of a potential rich source of evidence that needs careful consideration in terms of storage, archiving and auditing and retrieval.
ix. Staff should not be told what monitoring is happening except in exceptional circumstances.

1.12 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks

i. Fingerprinting
ii. Computer Analysis and Response Team (CART).
iii. Credibility

iv. need to know
v. International Association of Computer Investigative Specialists
vi. purpose
vii. metadata
viii. Intrusion Detection Systems.
ix. escalation
x. International Organization on Computer Evidence.

2. State true or false

i. True
ii. True
iii. True
iv. False
v. True
vi. False
vii. True
viii. True
ix. False

1.13 MODEL QUESTIONS

1. What are the four stages of computer forensic process?
2. What are the uses of computer forensics?
3. What are the objectives of computer forensics?
4. What is the role of a forensics investigator?
5. What is forensics readiness plan?
6. What are the benefits of forensic readiness?
7. What are various steps involved in forensic readiness planning?
8. What is continuity of evidence?

UNIT II: COMPUTER FORENSICS INVESTIGATION PROCESS

2.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand the process of investigating computer crime
- Apply initial decision making process
- Assess the situation
- Notify decision makers and acquire authorisation
- Review policies and laws related to forensics investigation process
- Acquire the data
- Analyse the data
- Report the investigation

2.2 INTRODUCTION TO COMPUTER CRIME INVESTIGATION

According to Warren G. Kruse II and Jay G. Heiser, authors of *Computer Forensics: Incident Response Essentials*, computer forensics is "the preservation, identification, extraction, documentation, and interpretation of computer media for evidentiary and/or root cause analysis." The computer investigation model shown in figure 1 organizes the different computer forensics elements into a logical flow[12].

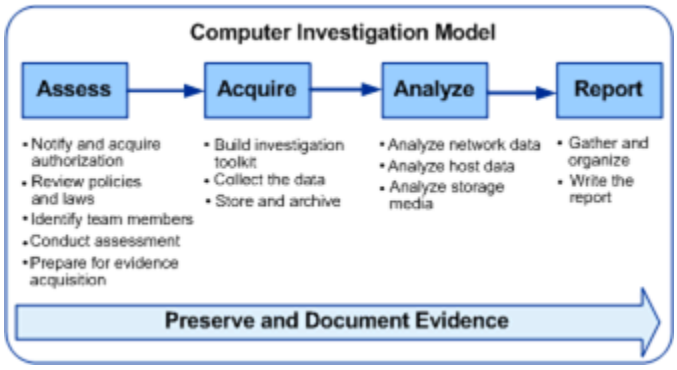

Figure 1: Computer investigation model

The four investigation phases and accompanying processes in the figure should be applied when working with digital evidence. The phases can be summarized as follows:

- **Assess the situation:** Analyze the scope of the investigation and the action to be taken.
- **Acquire the data:** Gather, protect, and preserve the original evidence.
- **Analyze the data:** Examine and correlate digital evidence with events of interest that will help you make a case.
- **Report the investigation:** Gather and organize collected information and write the final report.

[12] http://www.microsoft.com/en-us/download/details.aspx?id=23378

Detailed information about each of the phases is provided in the proceeding sections of this unit.

2.2.1 Initial Decision-Making Process

Before you begin each of the general investigation phases you should apply the initial decision-making process shown in the *Figure 2*.

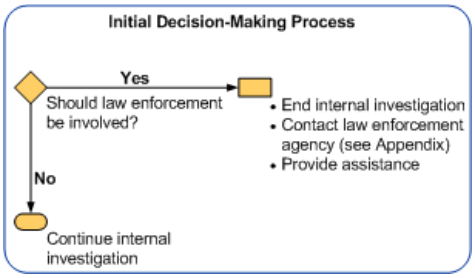

Figure 2: Initial decision making process

You should determine whether or not to involve law enforcement with the assistance of legal advisors. If you determine that law enforcement is needed, then you need to continue the internal investigation unless law enforcement officials advise you otherwise. Law enforcement might not be available to assist in the investigation of the incident, so you must continue to manage the incident and investigation for later submission to law enforcement.

Depending on the type of incident being investigated, the primary concern should be to prevent further damage to the organization by those person(s) who caused the incident. The investigation is important, but is secondary to protecting the organization unless there are national security issues.

2.3 ASSESS THE SITUATION

This section describes how to conduct a thorough assessment of the situation, how to establish scope, and the required resources for an internal investigation. Use the five-step process shown in the *Figure 3*.

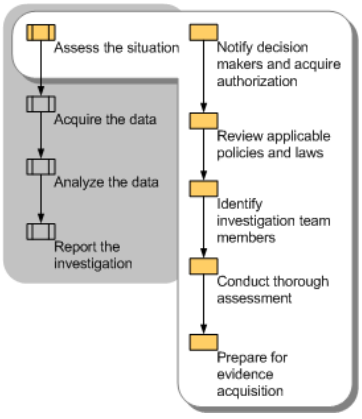

Figure 3: Assesment phase of the computer investigation model

2.3.1 Notify Decision Makers and Acquire Authorization

To conduct a computer investigation, you first need to obtain proper authorization unless existing policies and procedures provide incident response authorization. Then you need to conduct a thorough assessment of the situation and define a course of action. Use the following best practices:

- If no written incident response policies and procedures exist, notify decision makers and obtain written authorization from an authorized decision maker to conduct the computer investigation.
- Document all actions you undertake that are related to this investigation. Ensure there is a complete and accurate documented summary of the events and decisions that occurred during the incident and the incident response. This documentation may ultimately be used in court to determine the course of action that was followed during the investigation.
- Depending on the scope of the incident and absent any national security issues or life safety issues, the first priority is to protect the organization from further harm. After the organization is secure, restoration of services (if needed) and the investigation of the incident are the next priorities.

Decisions you make may be questioned as much as the evidence. Because computer evidence is complex, different investigations (such as those conducted by an opposing party) may make different decisions and reach different conclusions.

2.3.2 Review Policies and Laws

At the start of a computer investigation it is important to understand the laws that might apply to the investigation as well as any internal organization policies that might exist. Note the following important considerations and best practices:

- Determine if you have legal authority to conduct an investigation. Does your organization have policies and procedures that address the privacy rights of employees, contractors, or other persons using your network? Do any such policies and procedures specify the circumstances in which monitoring is allowed? Many organizations state in their policies and procedures that there is no expectation of privacy in the use of the organization's equipment, e-mail, Web services, telephone, or mail, and that the company reserves the right as a condition of employment to monitor and search these resources. Such policies and procedures should be reviewed by the organization's legal advisors, and all employees, contractors, and visitors should be notified of their existence. If you are uncertain about your authority, contact your management, your legal advisors, or (if necessary) your local authorities.
- Consult with your legal advisors to avoid potential issues from improper handling of the investigation. These issues may include:
    - Compromising customers' personal data.
    - Violating any state or federal law, such as federal privacy rules.
    - Incurring criminal or civil liability for improper interception of electronic communications. Consider warning banners.

    - Viewing sensitive or privileged information. Sensitive data that may compromise the confidentiality of customer information must only be made available as part of investigation-related documentation if it directly pertains to the investigation.
- Ensure the following customer privacy and confidentiality issues are addressed:
    - All data should be transferred securely, stored on local computers (not network servers), and should not be easily accessible.
    - All data (including documentation) should be maintained for the period specified by legal advisors or local policy after the computer investigation is closed. If the data is part of a potential criminal case, consult with the law enforcement agency investigating the case. If the case is a civil case, consult with your organization's legal advisors.
- Maintain digital copies of evidence, printouts of evidence, and the chain of custody for all evidence, in case of legal action. Preservation of the chain of custody is accomplished by having verifiable documentation that indicates who handled the evidence, when they handled it, and the locations, dates, and times of where the evidence was stored. Secure storage of evidence is necessary, or custody cannot be verified.

2.3.3 Identify Investigation Team Members

Determining who should respond to an incident is important to conducting a successful internal computer investigation. Ideally, team membership should be established before the team is needed for an actual investigation. It is important that investigation teams be structured appropriately and have appropriate skills. Your organization could establish team membership as part of a disaster recovery planning process. Use the following best practices as guidance for forming an investigation team:

- Identify a person who understands how to conduct an investigation. Remember that the credibility and skills of the person performing the investigation are often scrutinized if a situation results in legal proceedings in a court of law.
- Identify team members and clarify the responsibilities of each team member.
- Assign one team member as the technical lead for the investigation. The technical lead usually has strong technical skills and is experienced in computer investigations. In investigations that involve suspected parties who are technically skilled, you might need to select investigation team members who are more skilled than the suspected parties.
- Keep the investigation team as small as possible to ensure confidentiality and to protect your organization against unwanted information leaks.
- Engage a trusted external investigation team if your organization does not have personnel with the necessary skills.
- Ensure that every team member has the necessary clearance and authorization to conduct their assigned tasks. This consideration is especially important if any third-party personnel, such as consultants, are involved in the investigation.

Important The volatile nature of digital evidence makes it critical to conduct investigations in a timely manner. Be sure to secure availability of all team members for the duration of any investigation.

2.3.4 Conduct a Thorough Assessment

A thorough, clearly documented assessment of the situation is required to prioritize your actions and justify the resources for the internal investigation. This assessment should define the current and potential business impact of the incident, identify affected infrastructure, and obtain as thorough an understanding as possible of the situation. This information will help you define an appropriate course of action.

Use the following best practices to conduct a thorough assessment:

- Use all available information to describe the situation, its potential severity, potentially affected parties, and (if available) the suspected party or parties.
- Identify the impact and sensitivity of the investigation on your organization. For example, assess whether it involves customer data, financial details, health care records, or company confidential information. Remember to evaluate its potential impact on public relations. This assessment will likely be beyond the expertise of IT, and should be done in conjunction with management and legal advisors.
- Analyze the business impact of the incident throughout the investigation. List the number of hours required to recover from the incident, hours of downtime, cost of damaged equipment, loss of revenue, and value of trade secrets. Such an assessment should be realistic and not inflated. The actual costs of the incident will be determined at a later date.
- Analyze affected intangible resources, such as future impact on reputation, customer relationships, and employee morale. Do not inflate the severity of the incident. This analysis is for informational purposes only to help understand the scope of the incident. The actual impact will be determined at a later date. This assessment will likely be beyond the expertise of IT, and should be done in conjunction with management and legal advisors.

Use the following best practices to identify, analyze, and document the infrastructure and computers that are affected by the situation. Much of this guidance could have already been followed as part of a risk assessment process to prepare a disaster recovery plan.

- Identify the network(s) that are involved, the number of computers affected, and the type of computers affected.
- Obtain the network topology documentation, which should include a detailed network diagram that provides infrastructure information about servers, network hardware, firewalls, Internet connections, and other computers on the network.
- Identify external storage devices and any remote computers that should be included. External storage devices could include thumb drives, memory and flash cards, optical discs, and magnetic disks.
- Capture the network traffic over a period of time if live analysis is required. This type of analysis is only needed if you believe there is ongoing suspicious traffic on

the network, and is typically only performed after auditing and logging have been exhausted as sources of evidence.

Important Network sniffing (capturing network traffic) can be a breach of privacy, depending on the scope of the capture. You should therefore be very cautious about deploying network capture tools on your network.

- Use tools to examine the state of software applications and operating systems on computers that are likely affected. Useful tools for this task include the Windows application logs, system logs, and Windows Sysinternals PsTools.
- Examine affected file and application servers.

  Important Some of the information gathered during this assessment (such as running processes and data in memory) is captured by your tools in real time. You must ensure that any records or logs generated are securely stored to prevent losing this volatile data.

In addition, the following best practices can help you obtain a complete understanding of the situation.

- Build a timeline and map everything to it. A timeline is especially important for global incidents. Document any discrepancies between the date and time of hosts, such as desktop computers, and the system date and time.
- Identify and interview anyone who might be involved in the incident, such as system administrators and users. In some situations, such people might be external to the organization. Interviewing users and affected personnel often provides good results and insights into the situation. Interviews should be conducted by experienced interviewers.
- Document all interview outcomes. You will need to use them later to fully understand the situation.
- Retrieve information (logs) from internal and external facing network devices, such as firewalls and routers, which might be used in the possible attack path.
- Some information, such as IP address and domain name ownership, is often public by its nature. For example, you can use the *Whois* tool available at *https://www.whois.net/* and *https://www.arin.net/index.html* to identify an owner of an IP address.

2.3.5 Prepare for Evidence Acquisition

To prepare for the Acquire the Data phase, you should ensure that you have properly determined the actions and outcome of the Assess the Situation phase. A detailed document containing all information you consider relevant provides a starting point for the next phase and for the final report preparation. In addition, understand that if the incident becomes more than just an internal investigation and requires court proceedings, it is possible that all processes used in gathering evidence might be used by an independent third party to try and achieve the same results.

Such a document should provide detailed information about the situation and include the following:

- An initial estimate of the impact of the situation on the organization's business.
- A detailed network topology diagram that highlights affected computer systems and provides details about how those systems might be affected.
- Summaries of interviews with users and system administrators.
- Outcomes of any legal and third-party interactions.
- Reports and logs generated by tools used during the assessment phase.
- A proposed course of action.

Important Creating consistent, accurate, and detailed documentation throughout the computer investigation process will help with the ongoing investigation. This documentation is often critical to the project's success and should never be overlooked. As you create documentation, always be aware that it constitutes evidence that might be used in court proceedings. Before you begin the next phase, ensure that you have obtained a responsible decision maker's signoff on the documentation that you created during the assessment phase.

2.4 ACQUIRE THE DATA

This section discusses how to acquire the data that is necessary for the investigation. Some computer investigation data is fragile, highly volatile, and can be easily modified or damaged. Therefore, you need to ensure that the data is collected and preserved correctly prior to analysis. Use the three-step process shown in the following figure.

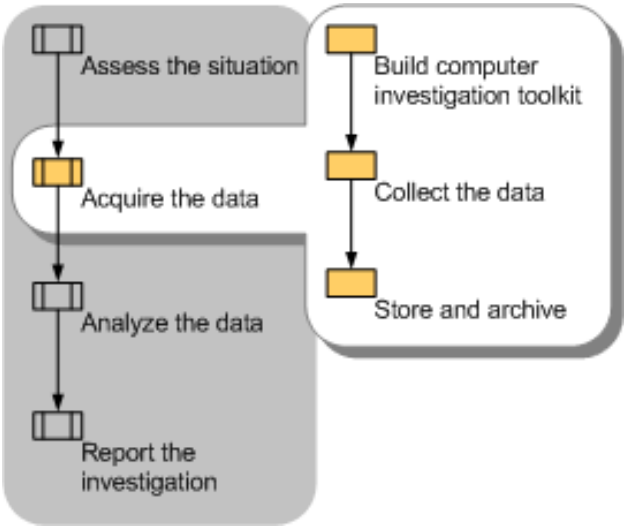

Figure 4: Acquisition phase of the computer investigation model

2.4.1 Build Computer Investigation Toolkit

Your organization will need a collection of hardware and software tools to acquire data during an investigation. Such a toolkit might contain a laptop computer with appropriate software tools, operating systems and patches, application media, write-protected backup devices, blank media, basic networking equipment, and cables. Ideally, such a toolkit will be created in advance, and team members will be familiar with the tools before they have to conduct an investigation.

2.4.1.1 Preparing Your Organization for a Computer Investigation

To prepare your organization for an internal computer investigation, you should assemble a readily available computer investigation toolkit that includes software and devices you can use to acquire evidence. Such a toolkit might contain a laptop computer with appropriate software tools, different operating systems and patches, application media, backup devices, blank media, basic networking equipment, and cables. Preparing this toolkit can be an ongoing task as you find the need for various tools and resources, depending upon the investigations you need to conduct.

Use the following guidelines when building and using a computer investigation toolkit:

- Decide which tools you plan to use before you start the investigation. The toolkit will typically include dedicated computer forensics software, such as Sysinternals, Encase, The Forensic Toolkit (FTK) , or ProDiscover.
- Ensure that you archive and preserve the tools. You might need a backup copy of the computer investigation tools and software that you use in the investigation to prove how you collected and analyzed data.
- List each operating system that you will likely examine, and ensure you have the necessary tools for examining each of them.
- Include a tool to collect and analyze metadata.
- Include a tool for creating bit-to-bit and logical copies.
- Include tools to collect and examine volatile data, such as the system state.
- Include a tool to generate checksums and digital signatures on files and other data, such as the File Checksum Integrity Validator (FCIV) tool.
- If you need to collect physical evidence, include a digital camera in the toolkit.

In addition, ensure that your toolkit meets the following criteria:

- Data acquisition tools are shown to be accurate. Proving accuracy is generally easier if you use well-known computer forensics software.
- The tools do not modify the access time of files.
- The examiner's storage device is forensically sterile, which means the disk drive does not contain any data, before it is used. You can determine whether a storage device is forensically sterile by running a checksum on the device. If the checksum returns all zeros, it does not contain any data.
- The examiner's hardware and tools are used only for the computer investigation process and not other tasks.

2.4.2 Collect the Data

Data collection of digital evidence can be performed either locally or over a network. Acquiring the data locally has the advantage of greater control over the computer(s) and data involved. However, it is not always feasible (for example, when computers are in locked rooms or other locations, or when high availability servers are involved). Other factors, such as the secrecy of the investigation, the nature of the evidence that must be gathered, and the timeframe for the investigation will ultimately determine whether the evidence is collected locally or over the network.

Important When using tools to collect data, it is important to first determine whether or not a rootkit has been installed. Rootkits are software components that take complete control of a computer and conceal their existence from standard diagnostic tools. Because rootkits operate at a very low hardware level, they can intercept and modify system calls. You cannot find a rootkit by searching for its executable, because the rootkit removes itself from the list of returned search results. Port scans do not reveal that the ports the rootkit uses are open, because the rootkit prevents the scanner from detecting the open port. Therefore, it is difficult to ensure that no rootkits exist.

When acquiring data over a network, you need to consider the type of data to be collected and the amount of effort to use. Consider what data you need to obtain that would support the prosecution of the offending parties. For example, it might be necessary to acquire data from several computers through different network connections, or it might be sufficient to copy a logical volume from just one computer.

The recommended data acquisition process is as follows:

1. Create accurate documentation that will later allow you to identify and authenticate the evidence you collect. Ensure that you note any items of potential interest and log any activities that might be of importance later in the investigation. Key to a successful investigation is proper documentation, including information such as the following:
   - Who performed the action and why they did it. What were they attempting to accomplish?
   - How they performed the action, including the tools they used and the procedures they followed.
   - When they performed the action (date and time) and the results.
2. Determine which investigation methods to use. Typically, a combination of offline and online investigations is used.
   - In offline investigations, additional analysis is performed on a bit-wise copy of the original evidence. (A bit-wise copy is a complete copy of all the data from the targeted source, including information such as the boot sector, partition, and unallocated disk space.) You should use the offline investigation method whenever possible because it mitigates the risk of damaging the original evidence. However, this method is only suitable for situations in which an image can be created, so it cannot be used to gather some volatile data.
   - In an online investigation, analysis is performed on the original live evidence. You should be especially careful when performing online analysis of data because of the risk of altering evidence that might be required to prove a case.
3. Identify and document potential sources of data, including the following:
   - Servers. Server information includes server role, logs (such as event logs), files, and applications.

- Logs from internal and external facing network devices, such as firewalls, routers, proxy servers, network access servers (NAS), and intrusion detection systems (IDS) that may be used in the possible attack path.
- Internal hardware components, such as network adapters (which include media access control (MAC) address information) and PCMCIA cards. Also note external port types, such as Firewire, USB, and PCMCIA.
- Storage devices that need to be acquired (internal and external), including hard disks, network storage devices, and removable media. Don't forget portable mobile devices such as PocketPC, Smartphone devices, and MP3 players such as Zune™.

4. When you must capture volatile data, carefully consider the order in which you collect the data. Volatile evidence can be easily destroyed. Information such as running processes, data loaded into memory, routing tables, and temporary files can be lost forever when the computer is shut down.
5. Use the following methods to collect data from storage media and record storage media configuration information:
    - If you need to remove any internal storage devices, turn off the computer first. However, before you turn off the computer you should verify that all volatile data has been captured whenever possible.
    - Determine whether to remove the storage device from the suspect computer and use your own system to acquire the data. It may not be possible to remove the storage device because of hardware considerations and incompatibilities. Typically, you would not disconnect storage devices such as RAID devices, storage devices with a hardware dependency (for example, legacy equipment), or devices in network storage systems such as storage area networks (SANs).
    - Create a bit-wise copy of the evidence in a backup destination, ensuring that the original data is write-protected. Subsequent data analysis should be performed on this copy and not on the original evidence. Step-by-step guidance for imaging is beyond the scope of this guide but is an integral part of evidence collection.

    Important Use industry accepted tools when acquiring a bit-wise copy. For example, EnCase FTK.

    - Document internal storage devices and ensure that you include information about their configurations. For example, note the manufacturer and model, jumper settings, and the size of the device. In addition, note the type of interface and the condition of the drive.
6. Verify the data you collect. Create checksums and digital signatures when possible to help establish that the copied data is identical to the original. In certain circumstances (for example, when a bad sector exists on the storage media) it may be impossible to create a perfect copy. Ensure that you have obtained the best copy possible with the available tools and resources. You can use the Microsoft File Checksum Integrity Verifier (FCIV) tool available at *http://www.microsoft.com/en-*

us/download/details.aspx?id=11533 to compute an MD5 or SHA1 cryptographic hash of the content of a file.

2.4.3 Store and Archive

When evidence is collected and ready for analysis, it is important to store and archive the evidence in a way that ensures its safety and integrity. You should follow any storage and archival procedures that exist within your organization.

Best practices for data storage and archival include the following:

- Physically secure and store the evidence in a tamperproof location.
- Ensure that no unauthorized personnel has access to the evidence, over the network or otherwise. Document who has physical and network access to the information.
- Protect storage equipment from magnetic fields. Use static control storage solutions to protect storage equipment from static electricity.
- Make at least two copies of the evidence you collected, and store one copy in a secure offsite location.
- Ensure that the evidence is physically secured (for example, by placing the evidence in a safe) as well as digitally secured (for example, by assigning a password to the storage media).
- Clearly document the chain of custody of the evidence. Create a check-in / check-out list that includes information such as the name of the person examining the evidence, the exact date and time they check out the evidence, and the exact date and time they return it.

2.5 ANALYZE THE DATA

This section discusses different approaches and well-accepted industry best practices for analyzing the evidence that is gathered during the Acquire the Data phase of an internal investigation. Use the three-step process shown in the following figure.

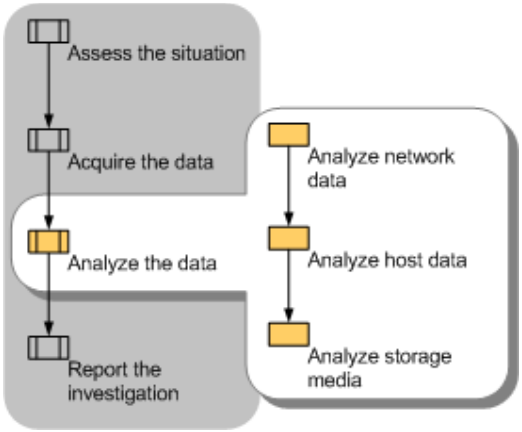

Figure 5: Analysis phase of the computer investigation model

Important Online analysis of data, which examines a computer directly while it is running, is often necessary. Online analysis is typically performed because of time constraints on an

investigation or to capture volatile data. You should be especially careful when performing online analysis to ensure that you minimize the risk to other evidence.

2.5.1 Analyze Network Data

In many investigations it is not necessary to analyze network data. Instead, the investigations focus on and examine images of the data. When network analysis is required, use the following procedure:

1. Examine network service logs for any events of interest. Typically, there will be large amounts of data, so you should focus on specific criteria for events of interest such as username, date and time, or the resource being accessed.
2. Examine firewall, proxy server, intrusion detection system (IDS), and remote access service logs. Many of these logs contain information from monitored incoming and outgoing connections and include identifying information, such as IP address, time of the event, and authentication information. You might want to examine the log data in a tool that is suited for data analysis, such as Microsoft® SQL Server™ 2005.
3. View any packet sniffer or network monitor logs for data that might help you determine the activities that took place over the network. In addition, determine whether connections you examine are encrypted—because you will not be able to read the contents of an encrypted session. However, you can still derive the time of the connection and whether a suspected party established a session with a specific server.

2.5.2 Analyze Host Data

Host data includes information about such components as the operating system and applications. Use the following procedure to analyze the copy of the host data you obtained in the Acquire the Data phase.

1. Identify what you are looking for. There will likely be a large amount of host data, and only a portion of that data might be relevant to the incident. Therefore, you should try to create search criteria for events of interest. For example, you might use the Microsoft Windows® Sysinternals Strings tool to search the files located in the \Windows\Prefetch folder. This folder contains information such as when and where applications were launched.
2. Examine the operating system data, including clock drift information, and any data loaded into the host computer's memory to see if you can determine whether any malicious applications or processes are running or scheduled to run. For example, you can use the Windows Sysinternals AutoRuns tool to show you what programs are configured to run during the boot process or login.
3. Examine the running applications, processes, and network connections. For example, you can look for running processes that might have an appropriate name but are running from non-standard locations.

2.5.3 Analyze Storage Media

The storage media you collected during the Acquire the Data phase will contain many files. You need to analyze these files to determine their relevance to the incident, which can be a

daunting task because storage media such as hard disks and backup tapes often contain hundreds of thousands of files.

Identify files that are likely to be relevant, which you can then analyze more closely. Use the following procedure to extract and analyze data from the storage media you collected:

1. Whenever possible, perform offline analysis on a bit-wise copy of the original evidence.
2. Determine whether data encryption was used, such as the Encrypting File System (EFS) in Microsoft Windows. Several registry keys can be examined to determine whether EFS was ever used on the computer. If you suspect data encryption was used, then you need to determine whether or not you can actually recover and read the encrypted data. Your ability to do so will depend upon different circumstances, such as the version of Windows, whether or not it is a domain-joined computer, and how EFS was deployed. For more information about EFS see "The Encrypting File System" on Microsoft TechNet. External EFS recovery tools are also available, such as Advanced EFS Data Recovery by Elcomsoft.
3. If necessary, uncompress any compressed files and archives. Although most forensic software can read compressed files from a disk image, you might need to uncompress archive files to examine all files on the media you are analyzing.
4. Create a diagram of the directory structure. It might be useful to graphically represent the structure of the directories and files on the storage media to effectively analyze the files.
5. Identify files of interest. If you know which files were affected by the security incident, you can focus the investigation on these files first. The hash sets created by the National Software Reference Library can be used to compare well-known files (such as operating system and application files) to the originals. Those files that match can normally be eliminated from the investigation. You can also use informational sites such as filespecs.com, Wotsit's Format, ProcessLibrary.com, and Microsoft DLL Help to help you categorize and collect information about existing file formats as well as to identify files.
6. Examine the registry, the database that contains Windows configuration information, for information about the computer boot process, installed applications (including those loaded during startup), and login information such as username and logon domain. For registry background information and detailed descriptions of registry content, see the Windows Server 2003 Resource Kit Registry Reference. Various tools are available for analyzing the registry, including RegEdit, which ships with the Windows operating system, Windows Sysinternals RegMon for Windows, and Registry Viewer by AccessData.
7. Search the contents of all gathered files to help identify files that may be of interest. Various intelligent searches can be performed using tools described in the "Tools" section in Appendix: Resources of this guide. For example, you can use the Windows Sysinternals Streams tool to reveal whether there are any NTFS alternate data streams used on files or folders. NTFS alternate data streams can

hide information within a file by causing it to appear to contain zero bytes of data when viewed through Windows Explorer although the file actually contains hidden data.

8. Study the metadata of files of interest, using tools such as Encase by Guidance Software, The Forensic Toolkit (FTK) by AccessData, or ProDiscover by Technology Pathways. File attributes such as timestamps can show the creation, last access, and last written times, which can often be helpful when investigating an incident.
9. Use file viewers to view the content of the identified files, which allow you to scan and preview certain files without the original application that created them. This approach protects files from accidental damage, and is often more cost effective than using the native application. Note that file viewers are specific to each type of file; if a viewer is not available, use the native application to examine the file.

After you analyze all of the available information, you may be able to reach a conclusion. However, it is important to be very cautious at this stage and ensure that you do not blame the wrong party for any damages. However, if you are certain of your findings, you will be ready to begin the Report the Investigation phase.

2.6 REPORT THE INVESTIGATION

This section discusses how to organize the information that you gather and the documentation that you create throughout a computer investigation, as well as how to write a final report. Use the two-step process shown in the following figure.

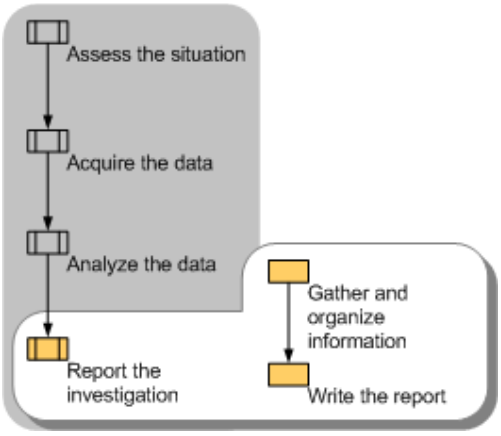

Figure 6: Reporting phase of the computer investigation model

2.6.1 Gather and Organize Information

During the initial phases of a computer investigation you create documentation about the specific activities in each phase. From within this documentation you need to identify the specific information that is relevant to your investigation and organize it into appropriate categories. Use the following procedure to gather and organize the required documentation for the final report.

1. Gather all documentation and notes from the Assess, Acquire, and Analyze phases. Include any appropriate background information.

2. Identify parts of the documentation that are relevant to the investigation.
3. Identify facts to support the conclusions you will make in the report.
4. Create a list of all evidence to be submitted with the report.
5. List any conclusions you wish to make in your report.
6. Organize and classify the information you gather to ensure that a clear and concise report is the result.

2.6.2 Write the Report

After you organize the information into appropriate categories, you can use it to write the final report. It is critical to the outcome of the investigation that the report is clear, concise, and written for the appropriate audience.

The following list identifies recommended report sections and information that should be included in these sections.

- **Purpose of Report**: Clearly explain the objective of the report, the target audience, and why the report was prepared.
- **Author of Report**: List all authors and co-authors of the report, including their positions, responsibilities during the investigation, and contact details.
- **Incident Summary**: Introduce the incident and explain its impact. The summary should be written so that a non-technical person such as a judge or jury would be able to understand what occurred and how it occurred.
- **Evidence**: Provide descriptions of the evidence that was acquired during the investigation. When describing evidence state how it was acquired, when, and who acquired it.
- **Details**: Provide a detailed description of what evidence was analyzed and the analysis methods that were used. Explain the findings of the analysis. List the procedures that were followed during the investigation and any analysis techniques that were used. Include proof of your findings, such as utility reports and log entries. Justify each conclusion that is drawn from the analysis. Label supporting documents, number each page, and refer to them by label name when they are discussed in the analysis. For example, "Firewall log from server, supporting document D." Also, provide information about those individuals who conducted or were involved with the investigation. If applicable, provide a list of witnesses.
- **Conclusion:** Summarize the outcome of the investigation. The conclusion should be specific to the outcome of the investigation. Cite specific evidence to prove the conclusion, but do not provide excessive detail about how the evidence was obtained (such information should be in the "Details" section). Include justification for your conclusion, along with supporting evidence and documentation. The conclusion should be as clear and unambiguous as possible. In many cases, it will be stated near the beginning of the report, because it represents the actionable information.
- **Supporting documents**: Include any background information referred to throughout the report, such as network diagrams, documents that describe the computer investigation procedures used, and overviews of technologies that are involved in the investigation. It is important that supporting documents provide enough information

for the report reader to understand the incident as completely as possible. As mentioned earlier, label each supporting document with letters and number each page of the document. Provide a complete list of supporting documents.

- If it is likely that the report will be presented to a varied audience, consider creating a glossary of terms used in the report. A glossary is especially valuable if the law enforcement agency is not knowledgeable about technical issues or when a judge or jury needs to review the documents.

2.7 SUMMARY

1. Computer forensics is "the preservation, identification, extraction, documentation, and interpretation of computer media for evidentiary and/or root cause analysis.
2. Depending on the type of incident being investigated, the primary concern should be to prevent further damage to the organization by those person(s) who caused the incident.
3. To conduct a computer investigation, you first need to obtain proper authorization unless existing policies and procedures provide incident response authorization.
4. At the start of a computer investigation it is important to understand the laws that might apply to the investigation as well as any internal organization policies that might exist.
5. Preservation of the chain of custody is accomplished by having verifiable documentation that indicates who handled the evidence, when they handled it, and the locations, dates, and times of where the evidence was stored.
6. Determining who should respond to an incident is important to conducting a successful internal computer investigation.
7. The volatile nature of digital evidence makes it critical to conduct investigations in a timely manner.
8. Creating consistent, accurate, and detailed documentation throughout the computer investigation process will help with the ongoing investigation.
9. Your organization will need a collection of hardware and software tools to acquire data during an investigation. Such a toolkit might contain a laptop computer with appropriate software tools, operating systems and patches, application media, write-protected backup devices, blank media, basic networking equipment, and cables.
10. Data collection of digital evidence can be performed either locally or over a network.
11. When using tools to collect data, it is important to first determine whether or not a rootkit has been installed.
12. When evidence is collected and ready for analysis, it is important to store and archive the evidence in a way that ensures its safety and integrity.
13. In many investigations it is not necessary to analyze network data. Instead, the investigations focus on and examine images of the data.
14. The storage media you collected during the Acquire the Data phase will contain many files.
15. After you organize the information into appropriate categories, you can use it to write the final report. It is critical to the outcome of the investigation that the report is clear, concise, and written for the appropriate audience.

2.8 CHECK YOUR PROGRESS

1. Fill in the blanks
   i. Assign one team member as the _____ for the investigation.
   ii. EFS stands for _____.
   iii. During the initial phases of a computer investigation you create _____about the specific activities in each phase.
   iv. If no written incident response policies and procedures exist, notify decision makers and obtain written authorization from an _____ decision maker to conduct the computer investigation.
   v. After the organization is secure, _____and the_____ of the incident are the next priorities.
   vi. Consult with your _____to avoid potential issues from improper handling of the investigation.
   vii. Preservation of the _____is accomplished by having verifiable documentation that indicates who handled the evidence, when they handled it, and the locations, dates, and times of where the evidence was stored.
   viii. Analyze the _____of the incident throughout the investigation.
   ix. Capture the _____ over a period of time if live analysis is required.
   x. _____ can be a breach of privacy, depending on the scope of the capture.
   xi. A_____ is especially important for global incidents.
   xii. _____users and affected personnel often provides good results and insights into the situation.
   xiii. As you create documentation, always be aware that it constitutes _____that might be used in court proceedings.
   xiv. _____ are software components that take complete control of a computer and conceal their existence from standard diagnostic tools.
   xv. Include a tool to collect and analyze _____.
2. State True or False
   i. The storage media you collected during the Acquire the Data phase will contain many files.
   ii. Inflate the severity of the incident.
   iii. Whenever possible, perform online analysis on a bit-wise copy of the original evidence.
   iv. Maintain digital copies of evidence, printouts of evidence, and the chain of custody for all evidence, in case of legal action.
   v. Engage a trusted external investigation team if your organization does not have personnel with the necessary skills.
   vi. Retrieve information (logs) from internal and external facing network devices, such as firewalls and routers, might be used in the possible attack path.

2.9 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks
   i. Technical lead
   ii. Encrypting File System.

iii. Documentation
iv. Authorized
v. restoration of services , investigation
vi. Legal advisors
vii. Chain of custody
viii. Business impact
ix. Network traffic
x. Network sniffing
xi. Timeline
xii. Evidence
xiii. Rootkits
xiv. Metadata

2. State true or false
    i. True
    ii. False
    iii. True
    iv. True
    v. True
    vi. True

2.10 MODEL QUESTIONS

1. What is computer forensics? Define.
2. What is network sniffing? List some popular tools used for packet sniffing.
3. What are the different phases of investigation process? Explain with the help of a diagram.
4. Why initial decision making process is important?
5. What are the different steps involved in the assessment of the situation?
6. What are the important guidelines for forming an investigating team?
7. What are the components of a computer investigation toolkit?
8. Explain the data acquisition process in detail.
9. List all the important sections that should be included in the investigation report.

UNIT III: DIGITAL EVIDENCE AND FIRST RESPONDER PROCEDURE

3.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Know about the digital evidence and best evidence rule
- Understand Locard's principle
- Identify various types of digital evidences
- Learn digital evidence investigation procedure
- Prepare first responder toolkit
- Create forensics tool testbed
- Document the forensics tool testbed and summary of the forensics tools
- Test the tools
- Recognise common mistakes of First Responder
- Identify various technical, administrative and legal issues of computer forensics
- Explain various types of investigations
- Classify techniques of digital forensics
- Understand volatile data
- Discover the importance of volatile data
- List order of volatility of digital evidences

3.2 DIGITAL EVIDENCE

Digital evidence[13] or **electronic evidence** is any probative information stored or transmitted in digital form that a party to a court case may use at trial. Before accepting digital evidence a court will determine if the evidence is relevant, whether it is authentic, if it is hearsay and whether a copy is acceptable or the original is required. Some of the popular electronic devices whaich are potential digital evidence are: HDD, CD/DVD media, backup tapes, USB drive, biometric scanner, digital camera, smart phone, smart card, PDA, etc.

The digital evidence are used to establish a crediable link between the attacker, victim, and the crime scene. Some of the information stored in the victim's system can be a potention digital evidence are IP address, system log-in & remote log-in details, browsing history, log files, emails, images, etc.

3.2.1 Locard's Principle

"Wherever a criminal steps, whatever he touches, whatever he leaves, even unconsciously, will serve as a silent witness against him. Not only his fingerprints or his footprints, but his hair, the fibers from his clothes, the glass he breaks, the tool mark he leaves, the paint he scratches, the blood or semen he deposits or collects. All of these and more, bear mute witness against him. This is evidence that does not forget. It is not confused by the excitement of the moment. It is not absent because human witnesses are. It is factual evidence. Physical

[13] https://en.wikipedia.org/wiki/Digital_evidence

evidence cannot be wrong, it cannot perjure itself, it cannot be wholly absent. Only human failure to find it, study and understand it, can diminish its value."

Digital evidence is usually not in a format that is directly readable by human. Therefore it requires some additional steps to convert it into a human readable form in the form of writing. Digital evidences must follow the requirements of the Best Evidence Rule.

3.2.2 Best Evidence Rule [14]

The best evidence rule, which had been established to deter any alteration of evidence, either intentionally or unintentionally, states that the court prefers the original evidence at the trial rather than a copy, but will accept a duplicate under these conditions:

- The original was lost or destroyed by fire, flood, or other acts of God. This has included such things as careless employees or cleaning staff.
- The original was destroyed in the normal course of business.
- The original is in possession of a third party who is beyond the court's subpoena power.

This rule has been relaxed to allow duplicates unless there is a genuine question as to the original's authenticity, or admission of the duplicate would, under the circumstances, be unfair.

3.2.3 Characteristics of Digital Evidence

Following are essential chaterchteristics of a digital evidence:

- **Admissibility:** It must be in confirmity with common law and legislative rules. There must be relationship between the evidence and the fact being proved. Digital evidence is often ruled inadmissible by courts because it was obtained without authorization. In most jurisdictions a warrant is required to seize and investigate digital devices. In a digital investigation this can present problems where, for example, evidence of other crimes are identified while investigating another.
- **Reliability**: The evidence must be from indisputed origin.
- **Completeness**: The evidence should prove the culprit 's actions and help to reach a conclusion.
- **Convincing to Judges:** The evidence must me convincing and understandable by the judges.
- **Authentication**: The evidence must be real and related to the incident. Courts largely concerned themselves with the reliability of such digital evidence. The investigator must be able to prove to the authenticity of the digital evidence by explaining:
    - the reliability of the computer equipment.
    - the manner in which the basic data was initially entered.
    - the measures taken to ensure the accuracy of the data as entered.
    - the method of storing the data and the precautions taken to prevent its loss.
    - the reliability of the computer programs used to process the data, and
    - the measures taken to verify the accuracy of the program.

[14] https://www.cccure.org/Documents/HISM/555-558.html

3.2.4 Stages in Digital Evidence Investigation Process[15]

Figure 7: Stages in digital evidence investigation process

- **Assesment:** It is a key point of an investigation where the potentially relevant sources of information are identified. Without this stage the chance to preserve and collect relevant material can be lost. This stage could also inform other activities including gathering information about possible passwords and attempts to attribute the sources to individuals as ownership of a device or a document can be a point of contention later on. In this phase the investigator make assesment of the situation anc consider many factor for making an assesment like whether the investigation is to be perforemed internally or an external agency is to be involved; Whether a search warrent is required. Also, some pre-search investigation need to be performed like gathering information about the infrastructure and assets of the company; gathering information about the employees who are directly or indirectly involved with the case; gathering information about the security incident team and their key skills, etc. Also the investigator need to prepare and check the forensic investigation toolkit to conduct the investigation. He also needs to brief the investigating team about the search strategy; guidelines to be followed while investigation for eg. maintaing the logs of the events, chain of evidence and chain of custody. Chain of evidence is the process of documenting each and every step carried out during the investigation process to prove the authenticity of the digital evidence in the court.
- **Acquisition:** It is a process of gathering the data from wherever it resides. The most common collection approach is to create an image of a target device which can then be examined without altering the original exhibit. In a wider sense, this could also apply to aspects such as requesting and receiving communications data. Cloud storage is an increasing concern and whilst the forensic recovery of files stored remotely is possible, the subsequent analysis may require detailed knowledge of the application used. Complications can also arise from the data being held in a different jurisdiction. The goal of the investigator in this phase is to acquire the evidence in a forensically sound manner so that it is acepted by court of law. It is good practice to record the physical attributes of every digital media like serial number, make, model, IP address and MAC address in case of network devices like NIC card, etc. and label them clearly so that they can easily be identified in later course of action. It is also a sound practice

[15] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/394779/ediscovery-digital-forensic-investigations-3214.pdf

to gather information regaring the user login, password, etc. from the users and system administrators. Remember to use forencially clean storage device to store the evidence. For making copy of the digital evicende, use bit-stream copy option, which acquire bit-by-bit image of the original evidence and can be considered as equalivent to original for the purpose of investigaion. Don't forget to calculate checksum or hash value of the on original copy and duplicate copy. The same value of checksum and hash value will gurentee both the copies are technically same for the purpose of investigation. It is important to note that logs from the servers, fairewalls, routers, and stand alone devices should also be recorded. Precautions regaring static electricty and magnetic fields should be taken while acquiring the digital evidence as it may alter data present in the digital devices. Therefore anti-static bags are used to store the digital evidence. The investigator must thoroughly examine the situation and if deemed essestial, a futher search warrent amy be required to search third party data carriers like ISP. After acqusation, the chain of custody, which the record of history of the custody of the evidence is prepared and recorded.

- **Preservation:** Preserving the digital evidence is as important as acquiring it and proper case must be taken to preserve the evidence so that data stored in digital storage devices can be used to investigate the case. It is advisable to take the photograph of the computer, cabling and the devices that are attached to the victim's computer, which are as important as victim's computer. Also label the seized cables along with the media. It is important to note that only forensically clean storage devices should be used to store the logs and other important digital information from the victim's system. Avoid dust, scratch, and exposure to magnetic or electric field by using antistatic bags. Care must also be taken to save the digital evidence from exposure to wireless radiations by storing them in wireless holdbags. One must avoid the use of USB drive or fireware drive as they change the state of the system. Intentional or accidental modification of data during the acquisition and preservation stage should be avoided and in case, it occurs, record the changes made in the system. Make the storage devices write-protect in order to accidentally overwrite the data. After bringing the digital media to the forensics lab, proper chain of custody should be maintain and these evidences should be stored in a physically safe location with access control facility. Also enough duplicate copies of the media should be made to carry out the investigation. *NEVER USE ORIGINAL MEDIA FOR CARRING OUT INVESTIGATION.*
- **Examination and Analysis:** The purpose of examination and analysis process is to make sense of the diverse digital data collected. A range of tools and techniques are used for this in an effort to ensure that as much data as possible is available for review. A lot of this data is of no relevance to the investigation but it may take considerable effort to get a good understanding of the relevance of material and to present it in an intelligible form. This data is examined and analysed to draw meaningful conclusions for the case. The first and the foremost thing to be kept in mind is the examination should be done be a trained person as mishandling of digital devices may corrupt the data. Examination requires the data to be extracted to the testbed for analysis. While examining, the goal of the investigator is to find out if files, folders, emails, partitions

are deleted and use recovery tools to restore them. Also, check if traces of data wiping software is present in the system so that special strategies could be use to recover data. If the files and documents are password protected then check whether the password for the same are available, else use password cracking software to crack the password and gain access to the files. The second important task after examination is analysis. It is the process of putting the different pieces of evidence together to allow conclusions to be drawn and ideas tested. Some units have dedicated analytical support available which is a useful resource but many investigators don not have routine access to analysts so it can be helpful for the investigator to be able to conduct their own analysis. The primary information is gathered based on the interviews conducted with the witnesses at the crime site which is then used to frame the keywords to search the relevant document, files, etc. for investigation. The photographs, paper documents seized during the raid, etc are useful for analysis. The Investigator look for document properties, file signatures, browser history, chat history, emails, printer spools, cache files, registry files, timeframe, ownership information, etc. to find clues and missing link. Hash values are compared to find weather a duplicate or multiple copies of the file exist. If required, use decrypting software to decrypt the files if they are encrypted. The most important point in the analysis process is to keep the log of all the steps carried out during the examination & analysis phase including the details of keywords used, the list of search results returned using these keywords, searching methodology used while carrying out investigation, etc.

- **Documentation and Reporting:** The examination and analysis can be conducted at a highly technical level but the information will ultimately need presenting to other individuals, either elsewhere in the investigation or the legal process, who are not so familiar with the detailed processes used and are more concerned with the usefulness of the information provided. Therefore documentation and reporting is a crucial part of the digital evidence investigation process. During this phase detailed report is perfored which includes all the information related to the case like details of OS, software, versions, patched installed in the machine and detailed note about the action taken during the forensic investigations along with the keywords seaches, logs, cache, etc . It also document any point that is contrary to the rules or to that which is normal or established. It also consists of the details of data analysing and the findings of the investigator.

3.3 FIRST RESPONDER TOOLKIT

The first responder is the person who first accesses the victim's computer. He must be prepared well to collect the evidences for the crime scene in a manner that is accepted by the court. Therefore, availability of trusted digital forensics toolkit is necessary for the first responder. Some of the important steps in preparing first responder's toolkit are:

1. Create forensics tool testbed.
2. Document the forensics tool testbed.
3. Document the summery of the forensics tools.
4. Test the tools.

The above four steps are described in details in the following section.

1. **Create forensic tool testbed-** The testbed should be created from the trusted source and functionality of the testbed should be checked in advance before using them in the field. Some of the guidelines are:
    a. Identify the appropriate OS type your organization is using, based on which the testbed is created. An organization may have verity of OS deployed in its network. For eg. it may have Linux based servers and Windows and Mac based PC/Laptop. In that case, one has to create multiple testbeds for each OS type.
    b. Disinfect the testbed from the availability of any data on the machine. Preferably use a new/fresh machine. In case, a new machine is not available use wiping tools to wipe out any data from the machine.
    c. Install OS and all the necessary software to conduct the forensics investigation.
    d. Ensure that the OS and all the programmes installed in the testbed are updated to latest version. If any patch is required for the successful operation of the the system, the same should also be installed.
    e. Compute Hash to ensure the integrity of the file system.
2. **Document the forensics tool testbed:** It includes the following
    a. Name, type and version of OS
    b. Details of the types of various applications/software installed in the testbed along with the details of the upgrades and patches.
    c. Details of various types of hardware installed in the testbed.
    d. Details pertaining to hash and checksum of the testbed.
3. **Document the summery of the forensics tools:** For every tool that is acquired for the testbed, the following information is documented for easy reference and record.
    a. Details about the source from where the software was brought. In case it's a freeware, mention the site/source from where the tool was downloaded.
    b. Detailed description about the purpose, working and compatibility of the tool with OS and other software.
    c. Details of tool dependencies and the system effects which include the details about the required system access levels by the user to run a tool and the details of shared libraries.
4. **Test the tools:** Now the tools selected and installed are tested in the testbed and its performance and output is examined.

3.3.1 Some Common Mistakes First Responder should avoid

- Do not shut-off or reboot the machine. This will erase all the valuable data present in the volatile devices.
- Do not assume that any parts of the victim/suspicious computer are reliable. Take precautions and follow procedures otherwise may accidently trigger malware which will effect/change/delete volatile data.

3.4 ISSUES FACING COMPUTER FORENSICS

The issues facing computer forensics examiners can be broken down into three broad categories: technical, legal and administrative[16].

[16] https://forensiccontrol.com/resources/beginners-guide-computer-forensics/

3.4.1 Technical issues

a. Encryption – Encrypted data can be impossible to view without the correct key or password. Examiners should consider that the key or password may be stored elsewhere on the computer or on another computer which the suspect has had access to. It could also reside in the volatile memory of a computer (known as RAM) which is usually lost on computer shut-down; another reason to consider using live acquisition techniques, as outlined above.

b. Increasing storage space – Storage media hold ever greater amounts of data, which for the examiner means that their analysis computers need to have sufficient processing power and available storage capacity to efficiently deal with searching and analysing large amounts of data.

c. New technologies – Computing is a continually evolving field, with new hardware, software and operating systems emerging constantly. No single computer forensic examiner can be an expert on all areas, though they may frequently be expected to analyse something which they haven't previously encountered. In order to deal with this situation, the examiner should be prepared and able to test and experiment with the behaviour of new technologies. Networking and sharing knowledge with other computer forensic examiners is very useful in this respect as it's likely someone else has already come across the same issue.

d. Anti-forensics – Anti-forensics is the practice of attempting to thwart computer forensic analysis. This may include encryption, the over-writing of data to make it unrecoverable, the modification of files' metadata and file obfuscation (disguising files). As with encryption, the evidence that such methods have been used may be stored elsewhere on the computer or on another computer which the suspect has had access to. In our experience, it is very rare to see anti-forensics tools used correctly and frequently enough to totally obscure either their presence or the presence of the evidence that they were used to hide.

3.4.2 Legal issues

Legal issues may confuse or distract from a computer examiner's findings. An example here would be the 'Trojan Defence'. A Trojan is a piece of computer code disguised as something benign but which carries a hidden and malicious purpose. A lawyer may be able to argue that actions on a computer were not carried out by a user but were automated by a Trojan without the user's knowledge; such a Trojan Defence has been successfully used even when no trace of a Trojan or other malicious code was found on the suspect's computer. In such cases, a competent opposing lawyer, supplied with evidence from a competent computer forensic analyst, should be able to dismiss such an argument. A good examiner will have identified and addressed possible arguments from the "opposition" while carrying out the analysis and in writing their report.

3.4.3 Administrative issues

a. Accepted standards – There are a plethora of standards and guidelines in computer forensics, few of which appear to be universally accepted. The reasons for this include: standard-setting bodies being tied to particular legislations; standards being aimed either at law enforcement or commercial forensics but not at both; the authors of such standards not being accepted by their peers; or high joining fees for professional bodies dissuading practitioners from participating.

b. Fit to practice – In many jurisdictions there is no qualifying body to check the competence and integrity of computer forensics professionals. In such cases anyone may present themselves as a computer forensic expert, which may result in computer forensic examinations of questionable quality and a negative view of the profession as a whole.

3.5 TYPES OF INVESTIGATION

There are four main types of investigation performed by digital forensics specialists[17]. The first three are broadly similar in the activities they involve, but differ in terms of the legal restrictions and guidelines imposed as well as the type of digital evidence and form of report.

3.5.1 Criminal forensics

The largest form of digital forensics and falling under the remit of law enforcement (or private contractors working for them). Criminal forensics is usually part of a wider investigation conducted by law enforcement and other specialists with reports being intended to facilitate that investigation and, ultimately, to be entered as expert evidence before the court. Focus is on forensically sound data extraction and producing report/evidence in simple terms that a lay man will understand.

3.5.2 Intelligence gathering

This type of investigation is often associated with crime, but in relation to providing intelligence to help track, stop or identify criminal activity. Unless the evidence is later to be used in court forensic soundness is less of a concern in this form of investigation, instead speed can be a common requirement.

3.5.3 Electronic discovery (eDiscovery)

Similar to "criminal forensics" but in relation to civil law. Although functionally identical to its criminal counter part, eDiscovery has specific legal limitations and restrictions, usually in relation to the scope of any investigation. Privacy laws (for example, the right of employees not to have personal conversation intercepted) and human rights legislation often affect electronic discovery.

3.5.4 Intrusion investigation

The final form of investigation is different from the previous three. Intrusion investigation is instigated as a response to a network intrusion, for example a hacker trying to steal corporate secrets. The investigation focuses on identifying the entry point for such attacks, the scope of access and mitigating the hackers activities. Intrusion investigation often occurs "live" (i.e. in real time) and leans heavily on the discipline of network forensics.

3.6 TECHNIQUES OF DIGITAL FORENSICS

A number of techniques are used during computer forensics investigations and much has been written on the many techniques used by law enforcement in particular[1].

3.6.1 Cross-drive analysis

[17] https://en.wikibooks.org/wiki/Introduction_to_Digital_Forensics/Types

A forensic technique that correlates information found on multiple hard drives. The process, still being researched, can be used to identify social networks and to perform anomaly detection.

3.6.2 Live analysis

The examination of computers from within the operating system using custom forensics or existing *sysadmin* tools to extract evidence. The practice is useful when dealing with Encrypting File Systems, for example, where the encryption keys may be collected and, in some instances, the logical hard drive volume may be imaged (known as a live acquisition) before the computer is shut down.

3.6.2.1 Volatile data

Volatile data is a data that is lost if the power is switched off. Computer requires some memory space where it could store most frequently used data, intermediately results of an operation, etc. which could be access by the CPU of a computer at faster rate. Some of the examples of fast memory are CPU registers, Cache memory, Random Access Memory(RAM), etc. The access time to these memory devices is low but they are volatile in nature. RAM contains wealth of information like system registries, passwords, browsing history, information about open processes and ports, uses profile of the system i.e. who logged into the computer, what are the hardware attached to the system, remote login details, IP address, etc. which could be very useful for the forensics investigator.

As discussed earlier, there are many volatile memory units present in system like CPU registers, Cache memory, RAM, etc. with different order of volatility. Order of volatility specifies the how sensitive the memory is towards the loss of data. Higher is the order of volatility, higher are the chances of data being lost/changed/modified. Therefore, the forensics investigator must follow the order of volatility to capture data from different memory devices. The order of volatility of various digital storage devices or digital evidences is shown in *Figure 8*. The higher is the level of memory in the pyramid, higher is the order of volatility.

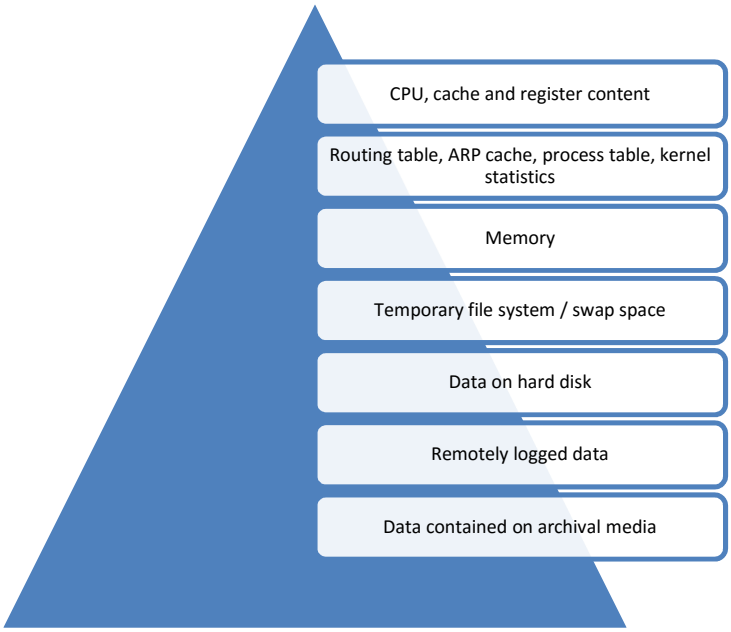
CPU, cache and register content
Routing table, ARP cache, process table, kernel statistics
Memory
Temporary file system / swap space
Data on hard disk
Remotely logged data
Data contained on archival media

Figure 8: Order of volatility of digital evidences

3.6.3 Recovery of Deleted files

A common technique used in computer forensics is the recovery of deleted files. Modern forensic software have their own tools for recovering or carving out deleted data. Most operating systems and file systems do not always erase physical file data, allowing investigators to reconstruct it from the physical disk sectors. File carving involves searching for known file headers within the disk image and reconstructing deleted materials.

3.6.4 Stochastic forensics

A method which uses stochastic properties of the computer system to investigate activities lacking digital artifacts. Its chief use is to investigate data theft.

3.6.5 Steganography

One of the techniques used to hide data is via steganography, the process of hiding data inside of a picture or digital image. An example would be to hide pornographic images of children or other information that a given criminal does not want to have discovered. Computer forensics professionals can fight this by looking at the hash of the file and comparing it to the original image (if available.) While the image appears exactly the same, the hash changes as the data changes. In Forensic examination, Steganalysis is used to get the details of Steganographic contents.

3.7 SUMMARY

1. Digital evidence or electronic evidence is any probative information stored or transmitted in digital form that a party to a court case may use at trial.

2. The digital evidences are used to establish a crediable link between the attacker, victim, and the crime scene.

3. Digital evidence is usually not in a format that is directly readable my human. Therefore it requires some additional steps to convert it into a human readable form in the form of writing.

4. There must be relationship between the evidence and the fact being proved.

5. The evidence must be from indisputed origin.

6. The evidence must be real and related to the incident.

7. Assessment is a key point of an investigation where the potentially relevant sources of information are identified.

8. Chain of evidence is the process of documenting each and every step carried out during the investigation process to prove the authenticity of the digital evidence in the court.

9. For making copy of the digital evicende, use bit-stream copy option, which acquire bit-by-bit image of the original evidence and can be considered as equalivent to original for the purpose of investigaion.

10. It is advisable to take the photograph of the computer, cabling and the devices that are attached to the victim's computer, which are as important as victim's computer.

11. Only forensically clean storage devices should be used to store the logs and other important digital information from the victim's system.

12. The examination of the digital evidence should be done by a trained person as mishandling of digital devices may corrupt the data.

13. The Investigator look for document properties, file signatures, browser history, chat history, emails, printer spools, cache files, registry files, timeframe, ownership information, etc. to find clues.

14. Do not shut-off or reboot the machine. This will erase all the valuable data present in the volatile devices.

15. A common technique used in computer forensics is the recovery of deleted files.

3.8 CHECK YOUR PROGRESS

1. Fill in the blanks

i. Digital evidences must follow the requirements of the _____.
ii. _____ are used to search the relevant files and documents from the digital evidence.
iii. _____ is the practice of attempting to thwart computer forensic analysis.
iv. A _____ is a piece of computer code disguised as something benign but which carries a hidden and malicious purpose.
v. _____ investigation is instigated as a response to a network intrusion.

vi. _____is a forensic technique that correlates information found on multiple hard drives.
vii. _____ analysis is useful when dealing with Encrypting File Systems.
viii. _____ data is a data that is lost of the power is switched off.
ix. _____ is the process of hiding data inside of a picture or digital image.

2. State True or False

i. Digital evidence is often ruled inadmissible by courts because it was obtained without authorization.
ii. The evidence must me convincing and it is not necessary that it should be understandable by the judges.
iii. Any storeage device can be used for storing the digital evidences.
iv. USB memory is a preferred way to store the logs and other information from the victim's computer.
v. Hash value of two duplicate file is same.
vi. Original media can be used to carry out digital investigation process.
vii. By default, every part of the victim's computer is considered unreliable.
viii. Encrypted data can be impossible to view without the correct key or password.
ix. Cache memory is an example of volatile memory.

3.9 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks

i. Best Evidence Rule.
ii. Keywords
iii. Antiforensics
iv. Trojan
v. Intrusion
vi. Cross-drive analysis
vii. Live
viii. Volatile
ix. Steganography

2. True or False

1. True
2. False
3. False
4. False
5. True
6. False
7. True
8. True
9. True

3.10 MODEL QUESTIONS

1. What is digital evidence? What is its role in the investigation process? Give examples of some common digital evidences.
2. State Locard's Principle.
3. What is best evidence rule? Under what circumstances the duplicate copy of the digital evidence is admissible for lawful purposes?
4. What are the essential characteristics of digital evidence?
5. How the authenticity of the digital evidence can be proved?
6. Explain the digital evidence investigation process in detail.
7. What is chain of evidence and chain of custody? Explain.
8. What is first responder's toolkit? What are the steps for preparing first responder's toolkit.
9. What are the various technical, legal and administrative issues faced by computer forensics?
10. What are the main types of investigation performed by digital forensics specialists?
11. What are the different techniques of digital forensics?
12. What is volatile data? What is order of volatility of digital evidences? Explain.

UNIT IV: UNDERSTANDING STORAGE MEDIA AND FILE SYSTEM

4.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Know about Hard Disk Drive(HDD)
- Understand the working of HDD
- Identify various types of interfaces
- Recognise internal structure of HDD
- List different type of formatting
- Understand booting process
- Discover the boot sequence of Windows, Mac and Linux OS
- Explain different type of file systems

4.2 HARD DISK DRIVE

A hard disk drive (HDD)[18], hard disk, hard drive or fixed disk is a data storage device used for storing and retrieving digital information using one or more rigid ("hard") rapidly rotating disks (platters) coated with magnetic material. The platters are paired with magnetic heads arranged on a moving actuator arm, which read and write data to the platter surfaces. Data is accessed in a random-access manner, meaning that individual blocks of data can be stored or retrieved in any order rather than sequentially. HDDs retain stored data even when powered off. The primary characteristics of an HDD are its capacity and performance. Capacity is specified in unit prefixes corresponding to powers of 1000: a 1-terabyte (TB) drive has a capacity of 1,000 gigabytes (GB; where 1 gigabyte = 1 billion bytes). Typically, some of an HDD's capacity is unavailable to the user because it is used by the file system and the computer operating system, and possibly inbuilt redundancy for error correction and recovery. Performance is specified by the time required to move the heads to a track or cylinder (average access time) plus the time it takes for the desired sector to move under the head (average latency, which is a function of the physical rotational speed in revolutions per minute), and finally the speed at which the data is transmitted (data rate).

4.2.1 Working

An HDD records data by magnetizing a thin film of ferromagnetic material on a disk. Sequential changes in the direction of magnetization represent binary data bits. The data is read from the disk by detecting the transitions in magnetization. User data is encoded using an encoding scheme, such as run-length limited encoding, which determines how the data is represented by the magnetic transitions.

A typical HDD design consists of a spindle that holds flat circular disks, also called platters, which hold the recorded data. The platters are made from a non-magnetic material, usually

[18] https://en.wikipedia.org/wiki/Hard_disk_drive

aluminum alloy, glass, or ceramic, and are coated with a shallow layer of magnetic material typically 10–20 nm in depth, with an outer layer of carbon for protection.

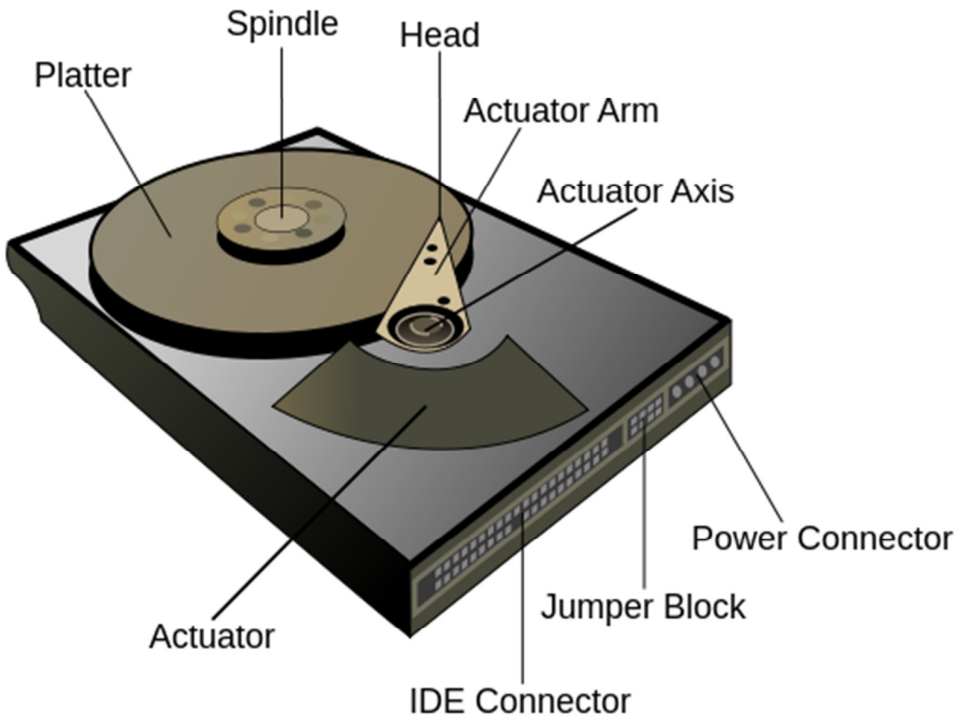

Figure 9: Hard disk drive of a computer[19]

The platters in contemporary HDDs are spun at speeds varying from 4,200 rpm in energy-efficient portable devices, to 15,000 rpm for high-performance servers. Information is written to and read from a platter as it rotates past devices called read-and-write heads that are positioned to operate very close to the magnetic surface, with their flying height often in the range of tens of nanometers. The read-and-write head is used to detect and modify the magnetization of the material passing immediately under it.

In modern drives there is one head for each magnetic platter surface on the spindle, mounted on a common arm. An actuator arm (or access arm) moves the heads on an arc (roughly radially) across the platters as they spin, allowing each head to access almost the entire surface of the platter as it spins. The arm is moved using a voice coil actuator or in some older designs a stepper motor. Early hard disk drives wrote data at some constant bits per second, resulting in all tracks having the same amount of data per track but modern drives (since the 1990s) use zone bit recording—increasing the write speed from inner to outer zone and thereby storing more data per track in the outer zones.

The two most common form factors for modern HDDs are 3.5-inch, for desktop computers, and 2.5-inch, primarily for laptops. HDDs are connected to systems by standard interface cables such as PATA (Parallel ATA), SATA (Serial ATA), USB or SAS (Serial attached SCSI) cables. The details of various types of HDD interfaces are discussed in the next section.

[19] https://en.wikipedia.org/wiki/Hard_disk_drive#/media/File:Hard_drive-en.svg

4.2.2 Interface

HDDs are accessed over one of a number of bus types, parallel ATA (PATA, also called IDE or EIDE; described before the introduction of SATA as ATA), Serial ATA (SATA), SCSI, Serial Attached SCSI (SAS), and Fibre Channel. Bridge circuitry is sometimes used to connect HDDs to buses with which they cannot communicate natively, such as IEEE 1394, USB and SCSI.

Modern interfaces connect an HDD to a host bus interface adapter with one data/control cable. Each drive also has an additional power cable, usually direct to the power supply unit. Now let us discuss various types of HDD interfaces in detail.

- **Small Computer System Interface (SCSI)**: originally named SASI for Shugart Associates System Interface, was standard on servers, workstations, Commodore Amiga, Atari ST and Apple Macintosh computers through the mid-1990s, by which time most models had been transitioned to IDE (and later, SATA) family disks. The range limitations of the data cable allows for external SCSI devices.

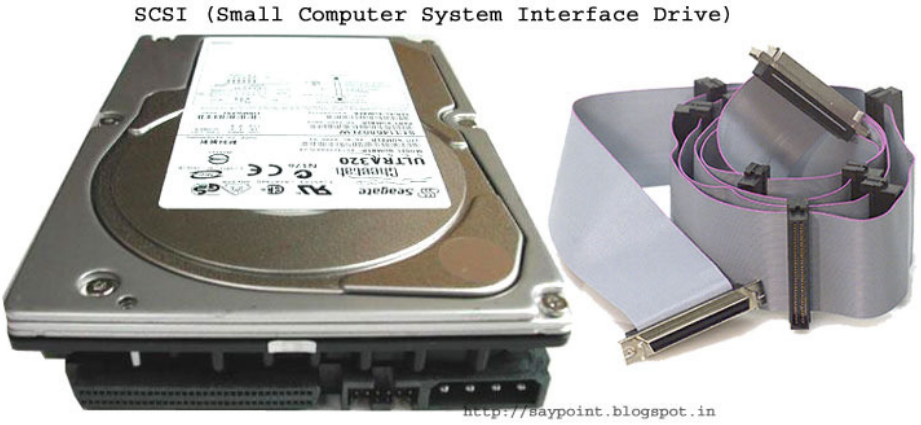

Figure 10: SCSI Interface[20]

- **Integrated Drive Electronics (IDE)**: later standardized under the name AT Attachment (ATA, with the alias P-ATA or PATA (Parallel ATA) retroactively added upon introduction of SATA) moved the HDD controller from the interface card to the disk drive. This helped to standardize the host/controller interface, reduce the programming complexity in the host device driver, and reduced system cost and complexity. The 40-pin IDE/ATA connection transfers 16 bits of data at a time on the data cable. The data cable was originally 40-conductor, but later higher speed requirements for data transfer to and from the HDD led to an "ultra DMA" mode, known as UDMA. Progressively swifter versions of this standard ultimately added the requirement for an 80-conductor variant of the same cable, where half of the conductors provide grounding necessary for enhanced high-speed signal quality by reducing cross talk.

[20] http://www.saypoint.net/2012/05/different-types-of-pc-hard-disk.html

Figure 11: Parallel ATA[21]

- **EIDE:** was an unofficial update (by Western Digital) to the original IDE standard, with the key improvement being the use of direct memory access (DMA) to transfer data between the disk and the computer without the involvement of the CPU, an improvement later adopted by the official ATA standards. By directly transferring data between memory and disk, DMA eliminates the need for the CPU to copy byte per byte, therefore allowing it to process other tasks while the data transfer occurs.
- **Fibre Channel (FC):** is a successor to parallel SCSI interface on enterprise market. It is a serial protocol. In disk drives usually the Fibre Channel Arbitrated Loop(FC-AL) connection topology is used. FC has much broader usage than mere disk interfaces, and it is the cornerstone of storage area networks (SANs). Recently other protocols for this field, like iSCSI and ATA over Ethernet have been developed as well. Confusingly, drives usually use *copper* twisted-pair cables for Fibre Channel, not fibre optics. The latter are traditionally reserved for larger devices, such as servers or disk array controllers.
- **Serial Attached SCSI (SAS)**: The SAS is a new generation serial communication protocol for devices designed to allow for much higher speed data transfers and is compatible with SATA. SAS uses a mechanically identical data and power connector to standard 3.5-inch SATA1/SATA2 HDDs, and many server-oriented SAS RAID controllers are also capable of addressing SATA HDDs. SAS uses serial communication instead of the parallel method found in traditional SCSI devices but still uses SCSI commands.
- **Serial ATA (SATA)**: The SATA data cable has one data pair for differential transmission of data to the device, and one pair for differential receiving from the device, just like EIA-422. That requires that data be transmitted serially. A similar differential signalling system is used in RS485, LocalTalk, USB, FireWire, and differential SCSI.

Figure 12: SATA[22]

[21] http://www.saypoint.net/2012/05/different-types-of-pc-hard-disk.html

4.3 DETAILS OF INTERNAL STRUCTURE OF HDD

Before[23] trying to understand formatting, you first need to understand how a hard drive works. Let's start with the detailed description of physical formatting (or low-level formatting) and logical formatting (or high-level formatting). Even though hard drives can be very small, they still contain millions of bits and therefore need to be organized so that information can be located. This is the purpose of the file system.

4.3.1 Low-Level Formatting

The purpose of low-level formatting is to divide the disk surface into basic elements:

- tracks
- sectors
- cylinders

Remember that a hard drive consists of several circular platters rotating around an axis and covered on either side by a magnetic oxide which, since it is polarised, can be used to store data.

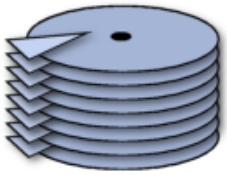

Figure 13: Platters in a HDD

The tracks are the concentric areas written on both sides of a platter.

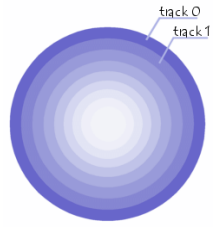

Figure 14: Tracks in a platter of HDD

Finally, these tracks are divided into pieces called sectors. There are millions of tracks and each has around 60 to 120 sectors.

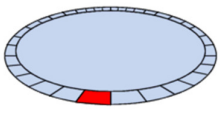

Figure 15: Tracks and sectors in a HDD

[22] http://www.saypoint.net/2012/05/different-types-of-pc-hard-disk.html
[23] http://ccm.net/contents/626-formatting-formatting-a-hard-drive

A **cylinder** refers to all the data located on the same track of different platters (i.e. vertically on top of each other) as this forms a "cylinder" of data in space.

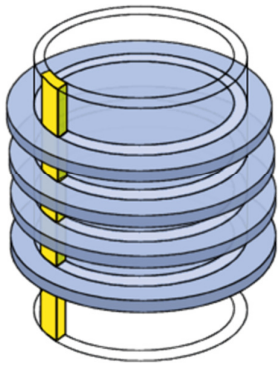

Figure 16: Cylinder in a HDD

Physical formatting therefore consists in organizing the surface of each platter into entities called trackers and sectors, by polarising the disk areas using the write heads. Tracks are numbered starting from 0, then the heads polarise concentrically the surface of the platters. When the head goes from one track to the next, it leaves a gap. Each track is itself organized into sectors (numbered starting from 1) and separated by *gaps*. Each of these sectors starts with an area reserved for system information called a *prefix* and ends with an area called a *suffix*. The purpose of low-level formatting is therefore to prepare the disk surface to receive data and to mark "defective sectors" using tests performed by the manufacturer. When you buy a hard drive, it has already undergone low-level formatting. SO YOU DO NOT NEED TO PERFORM LOW-LEVEL FORMATTING!

During the formatting, check tests (algorithms allowing the validity of sectors to be tested using checksums) are performed and each time a sector is considered defective, the (invalid) checksum is written in the prefix. It can no longer be used thereafter and is said to be "marked defective". When the disk reads the data, it sends a value that depends on the content of the sent packet, and which is initially stored with the data. The system calculates this value based on the data received, and then it compares it with the one that is stored with the data. If these two values are different, the data are no longer valid and there is probably a problem with the disk surface. The cyclic redundancy check (CRC), is based on the same principle to check the integrity of a file. Analysis utilities such as *scandisk* or *chkdsk* operate differently: they write data on sectors considered to be valid, and then read them and compare them. If they are the same, the utility goes on to the next sector, otherwise it marks the sector as defective.

4.3.2 High-level formatting

Logical formatting occurs after the low-level formatting. It creates a file system on the disks that will allow an operating system (DOS, Windows 95, Linux, OS/2, Windows NT, etc.) to use the disk space to store and access files. Operating systems use different file systems, so the type of logical formatting will depend on the operating system you install. So, if you format your disk with a single file system, this naturally limits the number and type of

operating systems that you can install (in fact, you can only install operating systems that use the same file system). Fortunately, there is a solution to this problem which is to create partitions. Each of the partitions can effectively have its own file system, and you can therefore install different types of operating systems.

What is a partition?

Partitioning is the process of writing the sectors that will make up the partition table (which contains information on the partition: size in sectors, position with respect to the primary partition, types of partitions present, operating systems installed, etc.). When a partition is created, it is given a *volume name* which allows it to be easily identified.

The partitioning of a hard drive occurs after the drive has been physically formatted but before it is logically formatted. It involves creating areas on the disk where data will not be mixed. It can be used, for example, to install different operating systems that do not use the same file system. There will therefore be at least as many partitions as there are operating systems using different file systems. If you are using just one operating system, a single partition the full size of the disk is sufficient, unless you want create several partitions so as to have, for example, several drives on which data are kept separate.

There are three types of partitions: **primary partitions**, **extended partitions** and **logical drives**. A disk may contain up to four primary partitions (only one of which can be active), or three primary partitions and one extended partition. In the extended partition, the user can create logical drives (i.e. "simulate" several smaller-sized hard drives).

Let's look at an example where the disk contains one primary partition and one extended partition made up of three logical drives (later we will look at multiple primary partitions):

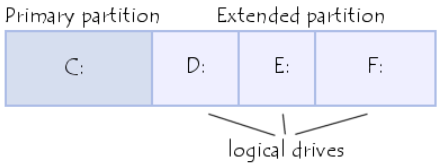

Figure 17: Partitining of HDD

For DOS systems (DOS, Windows 9x), only the primary partition is bootable, and is therefore the only one on which the operating system can be started.

4.3.3 Glossary of some important terms

4.3.3.1 Slack space

The unused space at the end of a file in a file system that uses fixed size clusters (so if the file is smaller than the fixed block size then the unused space is simply left). Often contains deleted information from previous uses of the block[24].

[24] https://en.wikipedia.org/wiki/Glossary_of_digital_forensics_terms

4.3.3.2 Lost Cluster

A lost cluster is a series of clusters on the hard disk drive that are not associated with a particular file. Data is there but it's unknown as to what file that data belongs to. Error checking tool looks for those and tries to repair them. Usually it can repair them, but sometimes it can't. If we're in a situation where we can't repair lost clusters that usually means we've got some serious disk errors and we're going to be losing some data[25].

4.3.3.3 Bad Sector

A bad sector is a sector on a computer's disk drive or flash memory that is either inaccessible or unwriteable due to permanent damage, such as physical damage to the disk surface or failed flash memory transistors[26].

4.3.3.4 Master Boot Record

The **boot sector**[27] (called the **Master Boot Record** or **MBR**) is the first sector of a hard drive (cylinder 0, head 0, sector 1), it contains the main partition table and the code, called the **boot loader**, which, when loaded into memory, will allow the system to boot up. After it is loaded into memory, this programme will determine from which system partition to boot, and will start the programme (called the *bootstrap*) which will start up the operating system present on that partition. This disk sector also contains all information concerning the hard drive (manufacturer, serial number, number of bytes per sector, number of sectors per cluster, number of sectors, etc.). This sector is therefore the most important one on the hard drive and is also used by the BIOS setup to recognize the hard drive. In other words, without it your hard drive is useless, which makes it a favourite target for viruses.

4.4 THE BOOTING PROCESS

The boot process[28] of a modern system involves multiple phases. The following components are involved in the boot process. They are each executed in this order:

- Power Supply Unit
- BIOS and CMOS
- POST tests
- Reading the Partition Table
- The Bootloader
- The Kernel
- OS Kernel

1. **Power Supply Unit:** When the power button is pressed, an electric circuit is closed which causes the power supply unit to perform a self test. In order for the boot process to continue, this self test has to complete successfully. If the power supply cannot confirm the self test, there will usually be no output at all. Most modern x86 computers, especially those using the ATX standard, will have two main connectors to the motherboard: a 4-pin connector to power the CPU, and a 24-pin connector to

[25] http://www.utilizewindows.com/pc-fundamentals/storage/333-hardware-and-software-disk-optimization
[26] https://en.wikipedia.org/wiki/Bad_sector
[27] http://ccm.net/contents/626-formatting-formatting-a-hard-drive
[28] http://ops-school.readthedocs.org/en/latest/boot_process_101.html

power other motherboard components. If the self test passes successfully, the PSU will send a signal to the CPU on the 4-pin connector to indicate that it should power on.

2. **BIOS and CMOS:** At its core, the Basic Input/Output System (*BIOS*) is an integrated circuit located on the computer's motherboard that can be programmed with firmware. This firmware facilitates the boot process so that an operating system can load.
   Let's examine each of these in more detail:
   a. *Firmware* is the software that is programmed into Electrically Erasable Programmable Read-Only Memory (EEPROM). In this case, the firmware facilitates booting an operating system and configuring basic hardware settings.
   b. An *integrated circuit* (IC) is what you would likely think of as a stereotypical "computer chip" - a thin wafer that is packaged and has metal traces sticking out from it that can be mounted onto a printed circuit board.

   Your BIOS is the lowest level interface you'll get to the hardware in your computer. The BIOS also performs the Power-On Self Test, or POST. Once the CPU has powered up, the first call made is to the BIOS. The first step then taken by the BIOS is to ensure that the minimum required hardware exists:

   - CPU
   - Memory
   - Video card

   Once the existence of the hardware has been confirmed, it must be configured. The BIOS has its own memory storage known as the CMOS (Complimentary Metal Oxide Semiconductor). The CMOS contains all of the settings the BIOS needs to save, such as the memory speed, CPU frequency multiplier, and the location and configuration of the hard drives and other devices.

   The BIOS first takes the memory frequency and attempts to set that on the memory controller. Next the BIOS multiply the memory frequency by the CPU frequency multiplier. This is the speed at which the CPU is set to run. Sometimes it is possible to "overclock" a CPU, by telling it to run at a higher multiplier than it was designed to, effectively making it run faster. There can be benefits and risks to doing this, including the potential for damaging your CPU.

3. **POST tests:** Once the memory and CPU frequencies have been set, the BIOS begins the Power-On Self Test (POST). The POST will perform basic checks on many system components, including:
   - Check that the memory is working
   - Check that hard drives and other devices are all responding
   - Check that the keyboard and mouse are connected (this check can usually be disabled)
   - Initialise any additional BIOSes which may be installed (e.g. RAID cards)
4. **Reading the Partition Table:** The next major function of the BIOS is to determine which device to use to start an operating system. A typical BIOS can read boot information from the devices below, and will boot from the first device that provides a successful response. The order of devices to scan can be set in the BIOS:

- Floppy disks
- CD-ROMs
- USB flash drives
- Hard drives
- A network

We'll cover the first four options here. For booting over the network, please refer to *http://networkboot.org/fundamentals/*.

There are two separate partition table formats: Master Boot Record (MBR) and the GUID Partition Table (GPT). We'll illustrate how both store data about what's on the drive, and how they're used to boot the operating system.

a. **Master Boot Record (the old way):** Once the BIOS have identified which drive it should attempt to boot from, it looks at the first sector on that drive. These sectors should contain the Master Boot Record.
   The MBR has two component parts:

   - The boot loader information block (448 bytes)
   - The partition table (64 bytes)

   The boot loader information block is where the first program the computer can run is stored. The partition table stores information about how the drive is logically laid out.

   The MBR has been heavily limited in its design, as it can only occupy the first 512 bytes of space on the drive (which is the size of one physical sector). This limits the tasks the boot loader program is able to do. The execution of the boot loader literally starts from the first byte. As the complexity of systems grew, it became necessary to add "chain boot loading". This allows the MBR to load an another program from elsewhere on the drive into memory. The new program is then executed and continues the boot process.

   If you're familiar with Windows, you may have seen drives labelled as "C:" and "D:" - these represent different logical "partitions" on the drive. These represent partitions defined in that 64-byte partition table.

b. **GPT - The GUID Partition Table (the new way):** The design of the IBM-Compatible BIOS is an old design and has limitations in today's world of hardware. To address this, the United Extensible Firmware Interface (UEFI) was created, along with GPT, a new partition format.
   There are a few advantages to the GPT format, specifically:

   - A Globally-Unique ID that references a partition, rather than a partition number. The MBR only has 64 bytes to store partition information - and each partition definition is 16 bytes. This design allows for unlimited partitions.

- The ability to boot from storage devices that are greater than 2 TBs, due to a larger address space to identify sectors on the disk. The MBR simply had no way to address disks greater than 2 TB.
- A backup copy of the table that can be used in the event that the primary copy is corrupted. This copy is stored at the 'end' of the disk.

There is some compatibility maintained to allow standard PCs that are using old BIOS to boot from a drive that has a GPT on it.

5. **The Bootloader:** The purpose of a bootloader is to load the initial kernel and supporting modules into memory.
6. **Kernel:** The kernel is the main component of any operating system. The kernel acts as the lowest-level intermediary between the hardware on your computer and the applications running on your computer. The kernel abstracts away such resource management tasks as memory and processor allocation. The kernel and other software can access peripherals such as disk drives by way of device drivers. Let us examine the booting process of some popular Operating Systems.

4.4.1 Linux Boot Process

The following are the 6 high level stages of a typical Linux boot process[29].

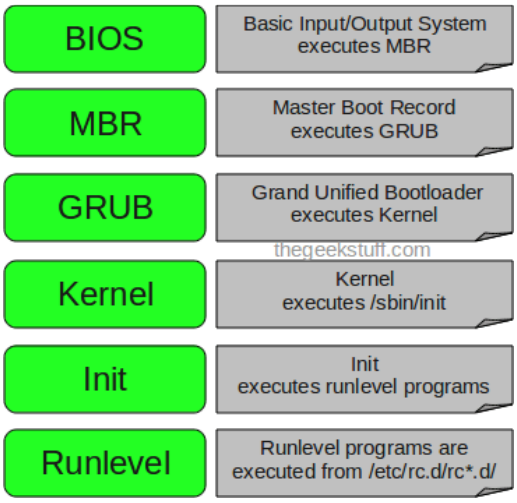

Figure 18: Linux boot process

Setp1: BIOS

- BIOS stands for Basic Input/Output System
- Performs some system integrity checks
- Searches, loads, and executes the boot loader program.
- It looks for boot loader in floppy, cd-rom, or hard drive. You can press a key (typically F12 of F2, but it depends on your system) during the BIOS startup to change the boot sequence.

[29] https://pacesettergraam.wordpress.com/2012/08/11/stages-of-linux-boot-process/

- Once the boot loader program is detected and loaded into the memory, BIOS gives the control to it.
- So, in simple terms BIOS loads and executes the MBR boot loader.

Step 2: MBR

- MBR stands for Master Boot Record.
- It is located in the 1st sector of the bootable disk. Typically /dev/hda, or /dev/sda
- MBR is less than 512 bytes in size. This has three components 1) primary boot loader info in 1st 446 bytes 2) partition table info in next 64 bytes 3) mbr validation check in last 2 bytes.
- It contains information about GRUB (or LILO in old systems).
- So, in simple terms MBR loads and executes the GRUB boot loader.

Step3: GRUB

- GRUB stands for Grand Unified Bootloader.
- If you have multiple kernel images installed on your system, you can choose which one to be executed.
- GRUB displays a splash screen, waits for few seconds, if you don't enter anything, it loads the default kernel image as specified in the grub configuration file.
- GRUB has the knowledge of the filesystem (the older Linux loader LILO didn't understand filesystem).
- Grub configuration file is /boot/grub/grub.conf (/etc/grub.conf is a link to this).
- GRUB just loads and executes Kernel and initrd images.

Step 4: Kernel

- Mounts the root file system as specified in the "root=" in grub.conf
- Kernel executes the /sbin/init program
- Since init was the 1st program to be executed by Linux Kernel, it has the process id (PID) of 1. Do a 'ps -ef | grep init' and check the pid.
- initrd stands for Initial RAM Disk.
- initrd is used by kernel as temporary root file system until kernel is booted and the real root file system is mounted. It also contains necessary drivers compiled inside, which helps it to access the hard drive partitions, and other hardware.

Step 5: Init

- Looks at the */etc/inittab* file to decide the Linux run level.
- Following are the available run levels
    - 0 – halt
    - 1 – Single user mode
    - 2 – Multiuser, without NFS
    - 3 – Full multiuser mode
    - 4 – unused

- 5 – X11
- 6 – reboot

- Init identifies the default initlevel from */etc/inittab* and uses that to load all appropriate program.
- Execute '*grep initdefault /etc/inittab'* on your system to identify the default run level
- If you want to get into trouble, you can set the default run level to 0 or 6. Since you know what 0 and 6 means, probably you might not do that.
- Typically you would set the default run level to either 3 or 5.

Step 6: Runlevel programs

- When the Linux system is booting up, you might see various services getting started. For example, it might say "starting sendmail …. OK". Those are the runlevel programs, executed from the run level directory as defined by your run level.
- Depending on your default init level setting, the system will execute the programs from one of the following directories.
  - Run level 0 – /etc/rc.d/rc0.d/
  - Run level 1 – /etc/rc.d/rc1.d/
  - Run level 2 – /etc/rc.d/rc2.d/
  - Run level 3 – /etc/rc.d/rc3.d/
  - Run level 4 – /etc/rc.d/rc4.d/
  - Run level 5 – /etc/rc.d/rc5.d/
  - Run level 6 – /etc/rc.d/rc6.d/
- Please note that there are also symbolic links available for these directory under /etc directly. So, /etc/rc0.d is linked to /etc/rc.d/rc0.d.
- Under the /etc/rc.d/rc*.d/ directories, you would see programs that start with S and K.
- Programs starts with S are used during startup. S for startup.
- Programs starts with K are used during shutdown. K for kill.
- There are numbers right next to S and K in the program names. Those are the sequence number in which the programs should be started or killed.
- For example, S12syslog is to start the syslog deamon, which has the sequence number of 12. S80sendmail is to start the sendmail daemon, which has the sequence number of 80. So, syslog program will be started before sendmail.

4.4.2 Mac OS Boot Sequence

The boot process starts with the activation of BootROM, the basic Macintosh ROM, which performs a Power On Self Test to test hardware essential to startup[30]. On the passing of this test, the startup chime is played and control of the computer is passed to OpenFirmware. OpenFirmware initializes the Random Access Memory, Memory Management Unit and hardware necessary for the ROM's operation. The OpenFirmware then checks settings, stored in NVRAM, and builds a list of all devices on a device tree by gathering their stored FCode information.

[30] https://en.wikipedia.org/wiki/BootX_(Apple)

On the completion of this task, BootX takes over the startup process configuring the keyboard and display, claiming and reserving memory for various purposes and checking to see if various key combinations are being pressed.After this process has been completed BootX displays the grey Apple logo, spins the spinning wait cursor, and proceeds to load the kernel and some kernel extensions and start the kernel.

The detailed description of the above is as follows:

- **BootROM:** As the name suggest, BootROM is a ROM (Read only Memory) which contains boot programmes viz. POST and Open Firmware.
    - POST: Power-On Self Test is the initial process which check the functionality of the basic hardware attached to the computer including RAM.
    - Open Firmware: The remaining hardware is initialized by Open Firmware. It also checks all the hardware associated with the systems and builds the initial device tree.
- **BootX:** The BootX initialize the kernel and the drivers required to boot the system from the cached set of device drivers. In case it is not present, it is loaded for /System/Library/Extensions for the same.
- **Kernel:** One the Kernel is loaded, it initialises the I/O kit which is used to control Input/Output devices. After this, the kernel initiates the launchd process.
- **Launchd:** it is process used for bootstrapping and is responsible for starting every system process. It also manages system initialization and starts the loginwindow process. During system initialization the system launchd process automatically starts /System/Library/LaunchDaemons, /Library/LaunchDaemons, /Library/StartupItems, and /etc/rc.local. The launch also manages daemons, a program who manages service request.
- **Startup Scripts and Startup Item:** As soon as /etc/rc.local is executed, it initialises the basic system by performing file-system consistency check and initiating SystemStarter process, a process used for launching startup items. For system configuration related information it refers to /etc/hostconfig.
- **The loginwindow Process:** this process displays the login screen that allows the user to authenticate, and then sets up and manages the graphical interface user environment based on the account preferences.
- **User Enviroment Setup:** After the user's credentials are authenticated, the user environment setup is performed based on the user's preference.

4.4.3 Boot Sequence in Windows 7

Figure 19: Window 7 boot sequence

1. In the first step, when the system is powered-on the *Basic Input Output System* (BIOS) and RAM is loaded and BIOS performs the hardware diagnostics based by initiating Power-On Self Test(POST).
2. In the second step, the BIOS locates MBR(Master Boot Record), which is located at the first sector of the first hard drive, to find the active drive, bootable partition and reads boot sector.
3. The boot sector loads *bootmgr*, which looks for the active partition on the drive and load *Boot Configuration Data(BCD)* data store. The information stored in BCD to find and load the selected operating system.
4. When the Windows 7 OS is selected, *bootmgr* executes a program called *winload.exe*, which takes the charge of the further process of loading windows. The following screen will appear in the monitor:

Figure 20: Windows login screen

5. The winload.exe starts ntoskrnl.exe which initiates all the necessary files needed to load and run Windows 7 OS.
6. The OS is loaded and *winlogon* is executed to provide the login interface. After authentication, the system settings are loaded based on the users credential and preferences.

4.5 FILE SYSTEM

Even though hard drives can be very small, they still contain millions of bits and therefore need to be organized so that information can be located[31]. This is the purpose of the file system. Remember that a hard drive is made up of several circular platters rotating around an axis. The tracks (concentric areas written to on either side of the platter) are divided into pieces called sectors (each 512 bytes in size). Logical formatting of a disk allows a file system to be created on the disk, which in turn will allow an operating system (DOS, Windows 9x, UNIX, etc) to use the disk space to store and use files. The file system is based on management of clusters, the smallest disk unit that the operating system is able to manage.

A cluster consists of one or more sectors, so the larger the cluster size, the fewer entities the operating system will have to manage. On the other hand, since an operating system only knows how to manage whole allocation units (i.e. a file occupies a whole number of clusters), the more sectors per cluster, the more wasted space there will be. This is why the choice of file system is important.

Files systems and the operating system: In reality, the choice of file system depends first of all on the operating system that you are using. In general, the more recent the operating system, the greater the number of files it will support. So, under DOS and on the first versions of Windows 95, FAT16 is required. Starting with Windows 95 OSR2, you have the choice between FAT16 and FAT32 file systems. If the partition size is greater than 2GB, then FAT file systems are excluded and you need to use the FAT32 system (or modify the size of the partition). Below this limit, FAT16 is recommended for partitions with a capacity of less than 500Mb, otherwise it is preferable to use FAT32.

In the case of Windows NT (up until version 4) you have the choice between the FAT16 system and NTFS, FAT32 is not supported. In general, the NTFS system is recommended as it provides higher security and better performance than the FAT system. Microsoft actually recommends using a small FAT-type partition (of between 250 and 500MB) for the operating system, so as to be able to boot from a bootable DOS floppy disk in case of a catastrophe, and to use a second partition for storing your data.

Under Windows NT5, there are more choices as it accepts FAT16, FAT32 and NTFS partitions. Once again, the more recent file system (NTFS 5) is recommended, as it offers many more features than the FAT systems. For the same reasons given above, you can still choose to have a FAT-type partition.

[31] http://ccm.net/contents/624-the-file-system

Table 2: Operating system and choice of file system

Operation system	File system types supported
Dos	FAT16
Windows 95	FAT16
Windows 95 OSR2	FAT16, FAT32
Windows 98	FAT16, FAT32
Windows NT4	FAT, NTFS (version 4)
Windows 2000/XP	FAT, FAT16, FAT32, NTFS (versions 4 and 5)
Linux	Ext2, Ext3, ReiserFS, Linux Swap(, FAT16, FAT32, NTFS)
MacOS	HFS (Hierarchical File System), MFS (Macintosh File System)
OS/2	HPFS (High Performance File System)
SGI IRIX	XFS
FreeBSD, OpenBSD	UFS (Unix File System)
Sun Solaris	UFS (Unix File System)
IBM AIX	JFS (Journaled File System)

Coexistence of several file systems: When several operating systems coexist on the same machine, the problem of choosing a file system is at its worse. Since the file system is tightly coupled to the operating system, when there are several operating systems you must choose a file system for each, bearing in mind that it is possible that data from one operating system may be accessed from another. One solution would be to use FAT partitions for all the systems, making sure that the partitions are no larger than 2 GB. The most appropriate solution would be to use for each OS a partition whose file system is best suited to it, and to use a a dedicated FAT16 partition for data to be shared by the different operating systems.

4.5.1 Some Common File systems

File systems are your interface to store your data[32]. Modern file systems offer a hierarchical view of your data, though historical file systems have been flat.

4.5.1.1 FAT

FAT stands for File Allocation Table, it is a relatively old file system, files are limited to 4GB in size and file names are case insensitive. It has the benefit of being widely portable, being available on many platforms. For this reason storage devices are often pre-formatted as FAT, just so less technical users don't assume the device is broken and return it.Its portability makes it useful for USB flash drives and the partition you use for /boot. It would be a poor choice for your root file system.

4.5.1.2 NTFS

This is Microsoft's primary file system. Its data structures don't limit the maximum file size to 4GB. As a commenter pointed out, NTFS is case sensitive, but not through the Windows API,

[32] https://yakking.branchable.com/posts/filesystems/

which maintains case insensitivity for compatibility with older software that assumed insensitivity, since it was dealing with FAT.
On Windows it is case preserving, so if you create a file called "Foo", you can read it as "foo", but when you list the contents of the directory, it is shown as "Foo", rather than "FOO", as FAT traditionally does. This makes it a better choice for storage media, as Linux is also able to read it. It is still inadvisable to use NTFS as your root file system on Linux, since it's primary use is reading disks that are also used by Windows machines, rather than being an installation's root file system.

4.5.1.3 ext2, ext3 and ext4

The ext file systems are Linux's primary file systems and are usually the default option when installing Linux distributions. Despite having a similar name, they are different beasts. ext2 is rather primitive, only really useful with old bootloaders. **ext3** is more advanced, though active development has moved on to ext4. **ext4** supports journalling, uses extents for its storage and supports extended attributes, where additional metadata can be assigned to a file. There are third-party tools to read ext file systems from Windows, but NTFS support in Linux is better than ext support in Windows.

4.5.1.4 XFS

XFS is a development from Silicon Graphics. It exceeds ext4's features, including the ability to take snapshots of the logical state of the file system, and was the source of extended attributes. It is available on IRIX and Linux, so portability is not its strong point, hence would not be useful on a USB flash drive, but it is an excellent choice for your root file system.

4.5.1.5 ZFS

ZFS is a product of Sun Microsystems, later bought by Oracle. It is a very advanced file system, offering all the features mentioned above plus more. This is a copy-on-write file system, unlike the above, which were either journalling, or wrote to the blocks directly. Being copy-on-write allows for deduplication, since if multiple files have the same data, the file system can point both files to the same data, since if it's changed in one file, the other one won't have its data changed. Live file system checking is possible with the scrub command, so downtime is not needed to perform maintenance. It can use multiple physical devices as its storage pool. Its license is incompatible with the Linux kernel, so kernel support is provided by a third-party module, which makes it possible that a kernel update could leave your file system unreadable, since the ZFS kernel module is un-readable. Loading an external kernel module is slower than it being built in, so this impacts boot speed. Despite its complexity, ZFS is also available on the Solaris and BSD unices.

4.5.1.6 BTRFS

Work on BTRFS was initiated by Oracle to be a competitor to ZFS, this is no longer the motivating factor, since Oracle acquired Sun, but BTRFS is likely to become the default Linux file system in the future. It is nearly as featureful as ZFS, only missing the online deduplication, which the BTRFS developers expect to complete in a couple of Linux kernel releases. Its design allows for a transition between ext and btrfs with the btrfs-convert tool, by saving the ext metadata elsewhere and re-mapping ext's data to btrfs extents. It still offers the original file system's data as a read-only disk image that can be mounted. Reverting back to ext is done by reinstating the ext metadata. Unfortunately, it has a reputation of being

unstable, corrupting your data and becoming unusable. This is a natural stage of file system maturity though, and BTRFS is my preferred root file system.

4.5.2 Types of file systems

File system types can be classified into disk/tape file systems, network file systems and special-purpose file systems[33].

4.5.2.1 Disk file systems

A *disk file system* takes advantages of the ability of disk storage media to randomly address data in a short amount of time. Additional considerations include the speed of accessing data following that initially requested and the anticipation that the following data may also be requested. This permits multiple users (or processes) access to various data on the disk without regard to the sequential location of the data. Examples include FAT (FAT12, FAT16, FAT32), exFAT, NTFS,HFS and HFS+, HPFS, UFS, ext2, ext 3, ext4, XFS, btrfs, ISO 9660, Files-11,Veritas File System, VMFS, ZFS, ReiserFS and UDF. Some disk file systems are journaling file systems or versioning file systems.

Optical discs: ISO 9660 and Universal Disk Format (UDF) are two common formats that target Compact Discs, DVDs and Blu-ray discs. Mount Rainier is an extension to UDF supported since 2.6 series of the Linux kernel and since Windows Vista that facilitates rewriting to DVDs.

4.5.2.2 Flash file systems

A *flash file system* considers the special abilities, performance and restrictions of flash memory devices. Frequently a disk file system can use a flash memory device as the underlying storage media but it is much better to use a file system specifically designed for a flash device.

4.5.2.3 Tape file systems

A *tape file system* is a file system and tape format designed to store files on tape in a self-describing form. Magnetic tapes are sequential storage media with significantly longer random data access times than disks, posing challenges to the creation and efficient management of a general-purpose file system.

In a disk file system there is typically a master file directory, and a map of used and free data regions. Any file additions, changes, or removals require updating the directory and the used/free maps. Random access to data regions is measured in milliseconds so this system works well for disks. Tape requires linear motion to wind and unwind potentially very long reels of media. This tape motion may take several seconds to several minutes to move the read/write head from one end of the tape to the other. Consequently, a master file directory and usage map can be extremely slow and inefficient with tape. Writing typically involves reading the block usage map to find free blocks for writing, updating the usage map and directory to add the data, and then advancing the tape to write the data in the correct spot. Each additional file write requires updating the map and directory and writing the data, which may take several seconds to occur for each file.

[33] https://en.wikipedia.org/wiki/File_system

Tape file systems instead typically allow for the file directory to be spread across the tape intermixed with the data, referred to as *streaming*, so that time-consuming and repeated tape motions are not required to write new data. However, a side effect of this design is that reading the file directory of a tape usually requires scanning the entire tape to read all the scattered directory entries. Most data archiving software that works with tape storage will store a local copy of the tape catalogue on a disk file system, so that adding files to a tape can be done quickly without having to rescan the tape media. The local tape catalogue copy is usually discarded if not used for a specified period of time, at which point the tape must be re-scanned if it is to be used in the future.

IBM has developed a file system for tape called the Linear Tape File System. The IBM implementation of this file system has been released as the open-source IBM Linear Tape File System — Single Drive Edition (LTFS-SDE) product. The Linear Tape File System uses a separate partition on the tape to record the index meta-data, thereby avoiding the problems associated with scattering directory entries across the entire tape.

Tape formatting

Writing data to a tape is often a significantly time-consuming process that may take several hours. Similarly, completely erasing or formatting a tape can also take several hours. With many data tape technologies it is not necessary to format the tape before over-writing new data to the tape. This is due to the inherently destructive nature of overwriting data on sequential media. Because of the time it can take to format a tape, typically tapes are pre-formatted so that the tape user does not need to spend time preparing each new tape for use. All that is usually necessary is to write an identifying media label to the tape before use, and even this can be automatically written by software when a new tape is used for the first time.

4.5.2.4 Database file systems

Another concept for file management is the idea of a database-based file system. Instead of, or in addition to, hierarchical structured management, files are identified by their characteristics, like type of file, topic, author, or similar rich metadata.

IBM DB2 for i (formerly known as DB2/400 and DB2 for i5/OS) is a database file system as part of the object based IBM i operating system (formerly known as OS/400 and i5/OS), incorporating a single level store and running on IBM Power Systems (formerly known as AS/400 and iSeries), designed by Frank G. Soltis IBM's former chief scientist for IBM i. Around 1978 to 1988 Frank G. Soltis and his team at IBM Rochester have successfully designed and applied technologies like the database file system where others like Microsoft later failed to accomplish. These technologies are informally known as 'Fortress Rochester' and were in few basic aspects extended from early Mainframe technologies but in many ways more advanced from a technological perspective.

Some other projects that aren't "pure" database file systems but that use some aspects of a database file system:

- Many Web content management systems use a relational DBMS to store and retrieve files. For example, XHTML files are stored as XML or text fields, while image files are

stored as blob fields; SQL SELECT (with optional XPath) statements retrieve the files, and allow the use of a sophisticated logic and more rich information associations than "usual file systems". Many CMSs also have the option of storing only metadata within the database, with the standard filesystem used to store the content of files.

- Very large file systems, embodied by applications like Apache Hadoop and Google File System, use some *database file system* concepts.

4.5.2.5 Transactional file systems

Some programs need to update multiple files "all at once". For example, a software installation may write program binaries, libraries, and configuration files. If the software installation fails, the program may be unusable. If the installation is upgrading a key system utility, such as the command shell, the entire system may be left in an unusable state.

Transaction processing introduces the isolation guarantee, which states that operations within a transaction are hidden from other threads on the system until the transaction commits, and that interfering operations on the system will be properly serialized with the transaction. Transactions also provide the atomicity guarantee, ensuring that operations inside of a transaction are either all committed or the transaction can be aborted and the system discards all of its partial results. This means that if there is a crash or power failure, after recovery, the stored state will be consistent. Either the software will be completely installed or the failed installation will be completely rolled back, but an unusable partial install will not be left on the system.

Windows, beginning with Vista, added transaction support to NTFS, in a feature called Transactional NTFS, but its use is now discouraged. There are a number of research prototypes of transactional file systems for UNIX systems, including the Valor file system, Amino, LFS, and a transactional ext3 file system on the TxOS kernel, as well as transactional file systems targeting embedded systems, such as TFFS.

Ensuring consistency across multiple file system operations is difficult, if not impossible, without file system transactions. File locking can be used as a concurrency control mechanism for individual files, but it typically does not protect the directory structure or file metadata. For instance, file locking cannot prevent TOCTTOU race conditions on symbolic links. File locking also cannot automatically roll back a failed operation, such as a software upgrade; this requires atomicity.

Journaling file systems are one technique used to introduce transaction-level consistency to file system structures. Journal transactions are not exposed to programs as part of the OS API; they are only used internally to ensure consistency at the granularity of a single system call.

Data backup systems typically do not provide support for direct backup of data stored in a transactional manner, which makes recovery of reliable and consistent data sets difficult. Most backup software simply notes what files have changed since a certain time, regardless of the transactional state shared across multiple files in the overall dataset. As a workaround, some database systems simply produce an archived state file containing all data up to that point, and the backup software only backs that up and does not interact directly with the active transactional databases at all. Recovery requires separate recreation of the database from the state file, after the file has been restored by the backup software.

4.5.2.6 Network file systems

A *network file system* is a file system that acts as a client for a remote file access protocol, providing access to files on a server. Programs using local interfaces can transparently create, manage and access hierarchical directories and files in remote network-connected computers. Examples of network file systems include clients for the NFS, AFS, SMB protocols, and file-system-like clients for FTP and WebDAV.

4.5.2.7 Shared disk file systems

A *shared disk file system* is one in which a number of machines (usually servers) all have access to the same external disk subsystem (usually a SAN). The file system arbitrates access to that subsystem, preventing write collisions. Examples include GFS2 from Red Hat, GPFS from IBM, SFS from DataPlow, CXFS from SGI and StorNext from Quantum Corporation.

4.5.2.8 Special file systems

A *special file system* presents non-file elements of an operating system as files so they can be acted on using file system APIs. This is most commonly done in Unix-like operating systems, but devices are given file names in some non-Unix-like operating systems as well.

Device file systems

A *device file system* represents I/O devices and pseudo-devices as files, called device files. Examples in Unix-like systems include devfs and, in Linux 2.6 systems, udev. In non-Unix-like systems, such as TOPS-10 and other operating systems influenced by it, where the full filename or pathname of a file can include a device prefix, devices other than those containing file systems are referred to by a device prefix specifying the device, without anything following it.

Other special file systems

- In the Linux kernel, *configfs* and *sysfs* provide files that can be used to query the kernel for information and configure entities in the kernel.
- *procfs* maps processes and, on Linux, other operating system structures into a filespace.

4.5.2.9 Minimal file system / audio-cassette storage

The late 1970s saw the development of the microcomputer. Disk and digital tape devices were too expensive for hobbyists. An inexpensive basic data storage system was devised that used common audio cassette tape. When the system needed to write data, the user was notified to press "RECORD" on the cassette recorder, then press "RETURN" on the keyboard to notify the system that the cassette recorder was recording. The system wrote a sound to provide time synchronization, then modulated sounds that encoded a prefix, the data, a checksum and a suffix. When the system needed to read data, the user was instructed to press "PLAY" on the cassette recorder. The system would *listen* to the sounds on the tape waiting until a burst of sound could be recognized as the synchronization. The system would then interpret subsequent sounds as data. When the data read was complete, the system would notify the user to press "STOP" on the cassette recorder. It was primitive, but it worked (a lot of the

time). Data was stored sequentially, usually in an unnamed format, although some systems (such as the Commodore PET series of computers) did allow the files to be named. Multiple sets of data could be written and located by fast-forwarding the tape and observing at the tape counter to find the approximate start of the next data region on the tape. The user might have to listen to the sounds to find the right spot to begin playing the next data region. Some implementations even included audible sounds interspersed with the data.

4.5.2.10 Flat file systems

In a flat file system, there are no subdirectories. When floppy disk media was first available this type of file system was adequate due to the relatively small amount of data space available. CP/M machines featured a flat file system, where files could be assigned to one of 16 *user areas* and generic file operations narrowed to work on one instead of defaulting to work on all of them. These user areas were no more than special attributes associated with the files, that is, it was not necessary to define specific quota for each of these areas and files could be added to groups for as long as there was still free storage space on the disk. The early Apple Macintosh also featured a flat file system, the Macintosh File System. It was unusual in that the file management program (Macintosh Finder) created the illusion of a partially hierarchical filing system on top of EMFS. This structure required every file to have a unique name, even if it appeared to be in a separate folder. While simple, flat file systems become awkward as the number of files grows and makes it difficult to organize data into related groups of files.

A recent addition to the flat file system family is Amazon's S3, a remote storage service, which is intentionally simplistic to allow users the ability to customize how their data is stored. The only constructs are buckets (imagine a disk drive of unlimited size) and objects (similar, but not identical to the standard concept of a file). Advanced file management is allowed by being able to use nearly any character (including '/') in the object's name, and the ability to select subsets of the bucket's content based on identical prefixes

4.6 SUMMARY

1. A hard disk drive (HDD), hard disk, hard drive or fixed disk is a data storage device used for storing and retrieving digital information using one or more rigid rapidly rotating disks (platters) coated with magnetic material.
2. Data is accessed in a random-access manner, meaning that individual blocks of data can be stored or retrieved in any order rather than sequentially. HDDs retain stored data even when powered off.
3. The primary characteristics of an HDD are its capacity and performance.
4. An HDD records data by magnetizing a thin film of ferromagnetic material on a disk.
5. A typical HDD design consists of a spindle that holds flat circular disks, also called platters, which hold the recorded data.
6. The two most common form factors for modern HDDs are 3.5-inch, for desktop computers, and 2.5-inch, primarily for laptops.
7. HDDs are connected to systems by standard interface cables such as PATA (Parallel ATA), SATA (Serial ATA), USB or SAS (Serial attached SCSI) cables.
8. The purpose of low-level formatting is to divide the disk surface into basic elements viz. tracks, sectors and cylinders.

9. Operating systems use different file systems, so the type of logical formatting will depend on the operating system you install.
10. The partitioning of a hard drive occurs after the drive has been physically formatted but before it is logically formatted.
11. There are three types of partitions: primary partitions, extended partitions and logical drives.
12. Partitioning is the process of writing the sectors that will make up the partition table.
13. The unused space at the end of a file in a file system that uses fixed size.
14. A lost cluster is a series of clusters on the hard disk drive that are not associated with a particular file.
15. A bad sector is a sector on a computer's disk drive or flash memory that is either inaccessible or unwriteable due to permanent damage, such as physical damage to the disk surface or failed flash memory transistors.
16. The boot sector is the first sector of a hard drive (cylinder 0, head 0, sector 1), it contains the main partition table and the code, called the boot loader, which, when loaded into memory, will allow the system to boot up.
17. Firmware is the software that is programmed into Electrically Erasable Programmable Read-Only Memory (EEPROM).
18. The purpose of a bootloader is to load the initial kernel and supporting modules into memory.
19. The kernel is the main component of any operating system. The kernel acts as the lowest-level intermediary between the hardware on your computer and the applications running on your computer.

4.7 CHECK YOUR PROGRESS

1. Fill in the blanks

i. The primary characteristics of an HDD are its _____ and _____.
ii. An HDD records data by magnetizing a thin film of _____ material on a disk.
iii. SCSI stands for _____.
iv. A _____ refers to all the data located on the same track of different platters.
v. _____ is the process of writing the sectors that will make up the partition table.
vi. BIOS stands for _____.
vii. _____is the software that is programmed into Electrically Erasable Programmable Read-Only Memory.
viii. CMOS stands for _____.
ix. The purpose of a _____is to load the initial kernel and supporting modules into memory.
x. _____is a development from Silicon Graphics.
xi. IBM has developed a file system for tape called the _____.
xii. A _____is a file system that acts as a client for a remote file access protocol, providing access to files on a server.

2. State True or False

i. The data is read from the disk by detecting the transitions in magnetization. True
ii. The platters are made from a non-magnetic material, usually aluminum alloy, glass, or ceramic.
iii. The 40-pin IDE/ATA connection transfers 32 bits of data at a time on the data cable.
iv. EIDE was an unofficial update (by Western Digital) to the original IDE standard.
v. EIDE have DMA transfer functionality.
vi. Physical level formatting is also known as high level formatting.
vii. The tracks are the concentric areas written on both sides of a platter.
viii. The partitioning of a hard drive occurs after the drive has been physically formatted but before it is logically formatted.
ix. The partition table stores information about how the drive is logically laid out.
x. The file system is based on management of clusters.
xi. The choice of file system does not depends on the operating system that you are using.
xii. NTFS system provides higher security and better performance than the FAT system.
xiii. Several operating systems coexist on the same machine.
xiv. The BIOS does not have its own memory storage.

4.8 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks

i. capacity , performance.
ii. ferromagnetic
iii. Small Computer System Interface.
iv. cylinder
v. Partitioning
vi. Basic Input/Output System.
vii. Firmware .
viii. Complimentary Metal Oxide Semiconductor.
ix. bootloader
x. XFS
xi. Linear Tape File System.
xii. network file system

2. State True or False

i. True
ii. True

iii. False
iv. True
v. True
vi. False
vii. True
viii. True
ix. True
x. True
xi. False
xii. True
xiii. True
xiv. False

4.9 MODEL QUESTIONS

1. What is Hard Disk Drive? What are its main characteristics?
2. Explain the working of HDD.
3. Explain various interfaces of HDD in detail.
4. Differentiate between high level formatting and low level formatting.
5. What is cyclic redundancy check (CRC)?
6. Define the terms:
    a. Slack space
    b. Lost cluster
    c. Bad sector
7. What is master boot record?
8. What is booting? Explain the booting process of Window 7 in detail.
9. What is a file system? Why it is used?
10. Give the details of file systems that different Operating System supports.
11. What is journaling?
12. What are different types of File Systems? Explain in detail.
13. What is flat file system?

References, Article Source & Contributors

[1]. *Bad Sector*. (2015, Sep. 26). Retrieved Oct. 21, 2015, from Wikipedia: https://en.wikipedia.org/wiki/Bad_sector available under the Creative Commons Attribution-ShareAlike License.

[2]. Boiarkine, V., Carter, R., Chappell, L., Cullimore, P., Quilty, T., Slater, P., et al. (2007, Jan.). *Fundamental Computer Investigation Guide for Windows.* (S. Wacker, Ed.) Retrieved Nov. 12, 2015, from http://www.microsoft.com/en-us/download/details.aspx?id=23378

[3]. *BootX (Apple)*. (2015, May 25). Retrieved Oct. 21, 2015, from Wikipedia: https://en.wikipedia.org/wiki/BootX_(Apple) available under the Creative Commons Attribution-ShareAlike License.

[4]. *Computer Storage*. (n.d.). Retrieved Sep. 21, 2015, from http://www.utilizewindows.com/: http://www.utilizewindows.com/pc-fundamentals/storage/333-hardware-and-software-disk-optimization available under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

[5]. Cooper, M. (2014, March). *General overview of the Linux file system*. Retrieved Oct. 22, 2015, from Linux Documentation Project: http://www.tldp.org/LDP/intro-linux/html/sect_03_01.html available under the terms of the GNU Free Documentation License.

[6]. Crye, A. B. (n.d.). *Locard's exchange principle explained*. Retrieved Jan. 19, 2016, from http://everything.explained.today/Locard's_exchange_principle/ available GNU Free Documentation License.

[7]. *Cryptography*. (n.d.). Retrieved Oct. 01, 2015, from ICSProjects.Wikispaces.com: http://icsproject.wikispaces.com/Cryptography available under a Creative Commons Attribution Share-Alike 3.0 License.

[8]. *CYBER SECURITY MANIFESTO 2.0*. (2012, Oct. 01). Retrieved Sep. 26, 2015, from cybersecuritymanifesto: http://cybersecuritymanifesto.com/

[9]. *Different Types Of PC Hard Disk Drives (HDD)*. (n.d.). Retrieved Oct. 21, 2015, from SAYPOINT: http://www.saypoint.net/2012/05/different-types-of-pc-hard-disk.html available under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

[10]. *Digital evidence*. (2015, Aug. 20). Retrieved Oct. 11, 2015, from Wikipedia: https://en.wikipedia.org/wiki/Digital_evidence available under the Creative Commons Attribution-ShareAlike License.

[11]. Edson, J. (2011, July 25). *A Brief History Of Forensic Science*. Retrieved Oct. 04, 2015, from riaus.org.au: http://riaus.org.au/articles/a-brief-history-of-forensic-science/ availableunder a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

[12]. ENISA, & Anderson, P. (2014). *Electronic evidence - a basic guide for First Responders under permission for reproduction of material published on the website https://www.enisa.europa.eu/about-enisa/legal-notice.* European Union Agency for Network and Information Security.

[13]. *File system*. (2015, Oct. 18). Retrieved Oct. 21, 2015, from Wikipedia: https://en.wikipedia.org/wiki/File_system available under the Creative Commons Attribution-ShareAlike License.

[14]. *File systems*. (2014, Jan.). Retrieved Oct. 21, 2015, from https://yakking.branchable.com/: https://yakking.branchable.com/posts/filesystems/ available under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

[15]. *File Systems: FAT, NTFS, and HFS+*. (n.d.). Retrieved Oct. 22, 2015, from www.study.com: http://study.com/academy/lesson/files-systems-fat-ntfs-hfs-and-ffs.html

[16]. *Formatting - Formatting a hard drive*. (2015, Oct. 01). Retrieved Oct. 21, 2015, from ccm.net: http://ccm.net/contents/626-formatting-formatting-a-hard-drive

[17]. *Formatting - Formatting a hard drive*. (2015, Sep.). Retrieved Oct. 21, 2015, from ccm.net: http://ccm.net/contents/626-formatting-formatting-a-hard-drive available under the Creative Commons Attribution - NonCommercial - ShareAlike 3.0 France license.

[18]. Gallagher, S. (2013, Oct. 02). *We are not who we are*. Retrieved Sep. 26, 2015, from Security Blog: https://securityblog.redhat.com/tag/two-factor-authentication/ available under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

[19]. Glass, E. (2003). *The NTLM Authentication Protocol and Security Support Provider*. Retrieved Sep. 26, 2015, from Sourceforge: http://davenport.sourceforge.net/ntlm.html Copyright © 2003, 2006 Eric Glass permission to use, copy, modify, and distribute

[20]. *Glossary of digital forensics terms*. (2015, Sep. 09). Retrieved Oct. 21, 2015, from Wikipedia: https://en.wikipedia.org/wiki/Glossary_of_digital_forensics_terms

[21]. (1998). How Email Works. In P. Grall, *How Internet Works* (p. 85). Que Corporation.

[22]. Gupta, A. (2011, March 01). *Digital Forensic Analysis Using BackTrack, Part 1*. Retrieved Sep. 26, 2015, from opensourceforu: http://opensourceforu.efytimes.com/2011/03/digital-forensic-analysis-using-backtrack-part-1/

[23]. Gupta, A. (2011, March 01). *Digital Forensic Analysis Using BackTrack, Part 1*. Retrieved Oct. 03, 2015, from OpenSourceForU: http://opensourceforu.efytimes.com/2011/03/digital-forensic-analysis-using-backtrack-part-1/ available under the creative commons Attribution-NonCommercial 3.0 Unported (CC BY-NC 3.0).

[24]. *Hard disk drive*. (2015, Oct. 16). Retrieved Oct. 21, 2015, from Wikipedia: https://en.wikipedia.org/wiki/Hard_disk_drive

[25]. Havercan, P. (2015, July 17). *A plain person's guide to Secure Sockets Layer*. Retrieved Sep. 26, 2015, from http://peter.havercan.net/computing/plain-persons-guide-to-secure-sockets-layer.html available under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

[26]. *How it works*. (2010, Jan. 17). Retrieved Sep. 26, 2015, from Wikidot: http://pychatter.wikidot.com/how-it-works available under Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.

[27]. *How Things Work/Hard Drive.* (2012, April 2012). Retrieved Oct. 22, 2015, from Wikibooks: https://en.wikibooks.org/wiki/Wikijunior:How_Things_Work/Hard_Drive available under the Creative Commons Attribution-ShareAlike License.

[28]. *How to Reveal a Fake Facebook Account*. (n.d.). Retrieved Sep. 27, 2015, from www.wikihow.com: http://www.wikihow.com/Reveal-a-Fake-Facebook-Account available Under an Attribution-Noncommercial-Share Alike 3.0 Creative Commons License.

[29]. *Introduction to computer forensics*. (n.d.). Retrieved Oct. 03, 2015, from Forensic Control: https://forensiccontrol.com/resources/beginners-guide-computer-forensics/

[30]. *Introduction to computer forensics*. (n.d.). Retrieved Oct. 11, 2015, from forensic control: https://forensiccontrol.com/resources/beginners-guide-computer-forensics/

[31]. *Introduction to computer forensics*. (n.d.). Retrieved Oct. 03, 2015, from Forensic Control: https://forensiccontrol.com/resources/beginners-guide-computer-forensics/ available undert the Creative Commons – Attribution Non-Commercial 3.0 license.

[32]. *Introduction to Cryptography*. (2001, Feb. 24). Retrieved Oct. 01, 2015, from efa.org.au: https://www.efa.org.au/Issues/Crypto/crypto1.html available under a Creative Commons Attribution 4.0 International Licence.

[33]. *Introduction to Digital Forensics*. (2011, Nov. 16). Retrieved Sep. 28, 2015, from Wikibooks: https://en.wikibooks.org/wiki/Introduction_to_Digital_Forensics

[34]. Kent, K., Chevalie, S., Grance, T., & Dang, H. (2006). *Guide to Integrating Forensic Techniques into Incident Response.* Special Publication 800-86, National Institute of Standard and Technology, U.S. Department of Commerce.

[35]. *Kerberos Authentication*. (n.d.). Retrieved Sep. 26, 2015, from Interactiva: http://computers.interactiva.org/Security/Authentication/Kerberos/ available under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

[36]. Lawton, D., Stacey, R., & Dodd, G. (2014). *eDiscovery in digital forensic investigation.* CAST publication number 32/14 available under the Open Government Licence v3.0 https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/.

[37]. Ligh, M. H., Case, A., Levy, J., & Walters, A. *The Art of Memory Forensics: Detecting Malware and Threats in Windows, Linux, and Mac Memory 1st Edition.* Wiley.

[38]. Mehnle, J. (2010, April 17). *Sender Policy Framework*. Retrieved Sep. 28, 2015, from Openspf: http://www.openspf.org/Introduction available under the Creative Commons CC BY-SA 2.5.

[39]. Morton, T. (2013, Sep. 13). *Types of investigations*. Retrieved Oct. 11, 2015, from Wikibooks: https://en.wikibooks.org/wiki/Introduction_to_Digital_Forensics/Types

[40]. Morton, T. (2013, Sep. 13). *Types of investigations*. Retrieved Oct. 04, 2015, from Introduction to Digital Forensics: https://en.wikibooks.org/wiki/Introduction_to_Digital_Forensics/Types available under the Creative Commons Attribution-ShareAlike License.

[41]. Mukasey, M. B., Sedgwick, J. L., & Hagy, D. W. (2008). *Electronic Crime Scene Investigation:A Guide for First Responders Second Edition.* Special report, National Institute of Justice .

[42]. *Password Authentication Protocol*. (2015, July 17). Retrieved Sep. 26, 2015, from WIKIPEDIA: https://en.wikipedia.org/wiki/Password_Authentication_Protocol available under the Creative Commons Attribution-ShareAlike License.

[43]. Peterson, D. (2015, July 06). *Computer Forensics Miami*. Retrieved Oct. 03, 2015, from computer-forensics.wikidot: http://computer-forensics.wikidot.com/ available under Creative Commons Attribution-ShareAlike 3.0 License.

[44]. *Public key cryptography*. (2005, Oct. 01). Retrieved Oct. 01, 2015, from iusmentis.com: http://www.iusmentis.com/technology/encryption/crashcourse/publickeycrypto/ available undert the creative commons Attribution-Share Alike 2.5 Netherlands (CC BY-SA 2.5 GB).

[45]. Quirk, S. (2014, Mar. 13). *Concordia Password Security Policy.* Retrieved Sep. 26, 2015, from http://kb.cu-portland.edu/Password+Security available under a Creative Commons Attribution 3.0 License.

[46]. *Recognise scam or hoax emails and websites*. (n.d.). Retrieved Sep. 27, 2015, from https://www.communications.gov.au: https://www.communications.gov.au/what-we-do/internet/stay-smart-online/your-identity/recognise-scam-or-hoax-emails-and-websites available under a Creative Commons

[47]. Rowlingson, R. (2005). *An Introduction to Forensic Readiness Planning.* available under Open Government Licence http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/, Centre for the Protection of National Infrastructure.

[48]. Saylor, A. (2015). *CRYPTOGRAPHY*. Retrieved Oct. 01, 2015, from learn.saylor.org: https://learn.saylor.org/course/cs409 available under a Creative Commons Attribution 3.0 Unported License.

[49]. Sbh. (2011, Jan. 14). *An Introduction to Cryptography*. Retrieved 10 01, 2015, from go4experts: http://www.go4expert.com/articles/introduction-cryptography-t24529/ available Under Creative Commons with Attribution license.

[50]. *Secret key cryptography*. (2005, Oct. 01). Retrieved Oct. 01, 2015, from iusmentis.com: http://www.iusmentis.com/technology/encryption/crashcourse/secretkeycrypto/ available under creative commons Attribution-Share Alike 2.5 Netherlands (CC BY-SA 2.5 GB) license.

[51]. *Selecting a strong password*. (2015, Sep. 10). Retrieved Sep. 26, 2015, from Wordpress: https://en.support.wordpress.com/selecting-a-strong-password/ available under Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

[52]. Smidsrød, R. (n.d.). *The fundamentals of network booting*. Retrieved Oct. 22, 2015, from Networkboot: http://networkboot.org/fundamentals/ available under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

[53]. *Stages of linux boot process*. (2012, Aug. 11). Retrieved Oct. 21, 2015, from https://pacesettergraam.wordpress.com: https://pacesettergraam.wordpress.com/2012/08/11/stages-of-linux-boot-process/ available under a Creative Commons Attribution-ShareAlike 4.0 International License.

[54]. Stewart, W. (2000, Jan. 07). *How Email Works*. Retrieved Sep. 28, 2015, from http://www.livinginternet.com/: http://www.livinginternet.com/e/ew.htm available under the Creative Commons Attribution Share Alike License.

[55]. Thakur, D. (n.d.). *What is Booting? Type of Booting*. Retrieved Oct. 22, 2015, from Computer Notes: http://ecomputernotes.com/fundamental/disk-operating-system/what-is-booting-type-of-booting

[56]. *The file system*. (2015). Retrieved Oct. 21, 2015, from ccm.net: http://ccm.net/contents/624-the-file-system available under the creative commons Attribution - NonCommercial - ShareAlike 3.0 France license.

[57]. The National Archives. (2011). *Digital Continuity to Support Forensic Readiness*. Retrieved Oct. 04, 2015, from nationalarchives: http://www.nationalarchives.gov.uk/documents/information-management/forensic-readiness.pdf available under the Open Government Licence v3.0.

[58]. *Understanding Authentication*. (2008, Feb. 14). Retrieved Sep. 26, 2015, from Go4Experts: http://www.go4expert.com/articles/understanding-authentication-t8842/ Under Creative Commons with Attribution license.

[59]. Verma, D. (2012, Nov. 05). *How To Identify Fake EMail And Trace Sender's Location*. Retrieved Sep. 27, 2015, from http://www.usethistip.com: http://www.usethistip.com/2012/11/how-to-identify-fake-email-and-trace.html available under a Creative Commons Attribution 3.0 License.

[60]. Vig, A. (2013). *The boot process*. Retrieved Oct. 21, 2015, from http://ops-school.readthedocs.org: http://ops-school.readthedocs.org/en/latest/boot_process_101.html available under Creative Commons license.

[61]. Wheelbarger, S. (2009, Aug. 27). *CyberForensics*. Retrieved Oct. 04, 2015, from Wikidot: http://colbycriminaljustice.wikidot.com/cyberforensics available under Creative Commons Attribution-ShareAlike 3.0 License.

[62]. Witjes, C., & Stremlau, A. (n.d.). *The Windows 7 Boot Process (sbsl)*. Retrieved Oct. 22, 2015, from http://social.technet.microsoft.com/: http://social.technet.microsoft.com/wiki/contents/articles/11341.the-windows-7-boot-process-sbsl.aspx

[63]. Wolfe, H. B. (2007). Electronic Forensics: A Case for First Responders. *19th Annual FIRST Conference on Computer Security Incident Handling.* Spain.

Further Reading

[1]. Bose, R. (2008). *Mcgraw Hill Education.* Mcgraw Hill Education.

[2]. Bunting, S., & Wei, W. (2006). *The Official EnCE: EnCase Certified ExaminorStudy Guide.* Wiley Publishing Inc.

[3]. Carrier, B. *File System Forensic Analysis.*

[4]. Carvey, H. *Windows Forensic Analysis DVD Toolkit, Second Edition.* SYNGRESS.

[5]. *Computer Forensics: Investigation Procedures and Response.* EC-Council Press.

[6]. Cowen, D. (2013). *Computer Forensics: A Beginners Guide.*

[7]. ENISA, & Anderson, P. (2014). *Electronic evidence - a basic guide for First Responders.* European Union Agency for Network and Information Security.

[8]. Gardner, M. (1972). *Codes, ciphers, and secret writing.*

[9]. Godbole, N., & Belapure, S. (2011). *Cyber Security (with CD): Understanding Cyber Crimes, Computer Forensics and Legal Perspectives.* Wiley.

[10]. (1998). How Email Works. In P. Grall, *How Internet Works* (p. 85). Que Corporation.

[11]. Hagen, W. V. (2002). *Linux Filesystems.* SAMS.

[12]. Kent, K., Chevalie, S., Grance, T., & Dang, H. (2006). *Guide to Integrating Forensic Techniques into Incident Response.* Special Publication 800-86, National Institute of Standard and Technology, U.S. Department of Commerce.

[13]. *Kerberos Authentication*. (n.d.). Retrieved Sep. 26, 2015, from Interactiva: http://computers.interactiva.org/Security/Authentication/Kerberos/

[14]. Krause, M., & Tipton, H. F. (Eds.). (1993). *Handbook of Information Security Management.* AUERBACH.

[15]. Lawton, D., Stacey, R., & Dodd, G. (2014). *eDiscovery in digital forensic investigation.* CAST publication number 32/14 available under the Open Government Licence v3.0 https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/.

[16]. Ligh, M. H., Case, A., Levy, J., & Walters, A. *The Art of Memory Forensics: Detecting Malware and Threats in Windows, Linux, and Mac Memory 1st Edition.* Wiley.

[17]. Mukasey, M. B., Sedgwick, J. L., & Hagy, D. W. (2008). *Electronic Crime Scene Investigation:A Guide for First Responders Second Edition.* Special report, National Institute of Justice .

[18]. Nelson. (2013). *Guide to Computer Forensics and Investigations.*

[19]. Nelson, B., Phillips, A., & Steuart, C. (2009). *Guide to Computer Forensics and Investigations.* Cengage Learning.

[20]. Nolan, R., O'Sullivan, C., Branson, J., & Waits, C. (2005). *First Responders guide to Computer Forensic.* CERT Training and Education.

[21]. Pachghare, V. K. (2008). *Cryptography and Information Security: Second Edition.* PHI.

[22]. Schneier, B. (1994). *Applied Cryptography.* Wiley.

[23]. Silberschatz, Galvin, & Gagne. (2006). *Operating System Principles.* Wiley.

[24]. Singh, A. (2006). *Mac OS X Internals: A Systems Approach.*

[25]. Singh, S. (1999). *The Code Book: The Evolution Of Secrecy From Mary, Queen Of Scots To Quantum Cryptography.* New york: Doubleday.

[26]. Vacca, J. (2009). *Computer Forensics: Computer Crime Scene Investigation.*

[27]. *Windows Disk File Systems.* (2010). General Books LLC.

[28]. Witjes, C., & Stremlau, A. (n.d.). *The Windows 7 Boot Process (sbsl)*. Retrieved Oct. 22, 2015, from http://social.technet.microsoft.com/: http://social.technet.microsoft.com/wiki/contents/articles/11341.the-windows-7-boot-process-sbsl.aspx

[29]. Wolfe, H. B. (2007). Electronic Forensics: A Case for First Responders. *19th Annual FIRST Conference on Computer Security Incident Handling.* Spain.

BLOCK II

UNIT I: WINDOWS FORENSICS

1.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand and appreciate the need for windows forensics.
- Understand various technical terminologies associated to forensics in windows systems.
- Identify major components and aspects of windows which are relevant during forensics.
- Understand basic technologies and tools used to carry out data capture from a windows system during forensic investigation.
- Understand basic tools and technologies behind capturing registry information from windows systems during forensic investigation.
- Understand technologies and tools behind data and file recovery in windows system during forensic investigation.

1.2 INTRODUCTION

Computer forensics involves analysis of a computer system and identifies traces or evidences of activities leading to a criminal activity. In a sense much of the criminal activities in current world have more than one link to computing environments or at least has some or other relation to computers. Most of the criminal/other investigation tends to find traces of data or information in a computer system that can lead to conclusion or at least leads to support a theory pertaining a criminal offence. Windows forensics involves analysing various aspects of windows for malicious or suspicious traces of data in order to reach an evidential conclusion of any case. Windows forensics process is to analyse gathered information from activities that took place in a windows system. Aspects of windows like the registry, files, cookies, bins, memory status etc. contains initial information that can be used to promise a conclusion.

1.2.1 Background and need for Window forensics

Among the major operating system in use, Microsoft window is the most widely used operating system. The Microsoft windows versions that are currently in use are; Windows 8 and Windows 10. Microsoft Windows originated in 1985, as an operating environment running on top of MS DOS, which was the standard operating system shipped on most of Intel architecture PCs.

In 1995, Windows 95 was released which only used MS-DOS as a bootstrap. For backwards compatibility, Win9x could run real-mode MS-DOS and 16 bits Windows3.x drivers. Windows ME, released in 2000, was the last version in the Win9x family. Later versions have all been based on the Windows NT kernel. Server editions of Windows are widely used. In recent years, Microsoft has expended significant capital in an effort to promote the use of Windows as a server operating system. However, Windows' usage on servers is not as widespread as on personal computers

To know about windows artefacts is quite important for digital forensics examiners, almost 90 percent of traffic in networks comes from computers using Windows as their operating system and the investigators will be most likely to encounter Windows and have to collect evidence from it in most of the cybercrime cases. Below, we will discuss several places from which evidence may be gathered and ways to collect information from Windows.

This chapter focuses on Windows forensics. It starts by covering the different types of volatile and non-volatile information an investigator can collect from a Windows system. It then goes into detail about collecting and analysing data in memory, the registry, and files.

1.2.2 Major forensic areas in windows

More generally an investigator likes to access and analyse following areas in windows:

a) Volatile information like, system time, logged users, open files, network information and drives that are mapped shared folders etc. These and many more aspects will be discovered in the next section under the windows volatile information head.
b) Non-volatile information like file systems, registry settings, logs, devices, slack space, swap file, indexes, partitions etc. these and many more will be discovered in coming section under the heading non-volatile information.
c) Windows memory like memory dumps and analysing dumps and other aspects.
d) Caches, cookies and history analysis.
e) Other aspects like recycle bins, documents, short cut files, graphics file, executable files etc.

1.2.2.1 Volatile information

Volatile Information can disappear or be easily modified. It retains its contents while powered on but when the power is interrupted the stored data is immediately lost. Following are few methods/tools to acquire some volatile information in a Windows system.

To get history of commands used on the computer we can use Doskey. Doskey is a utility for DOS and Microsoft Windows that adds command history (see figure 2.1).

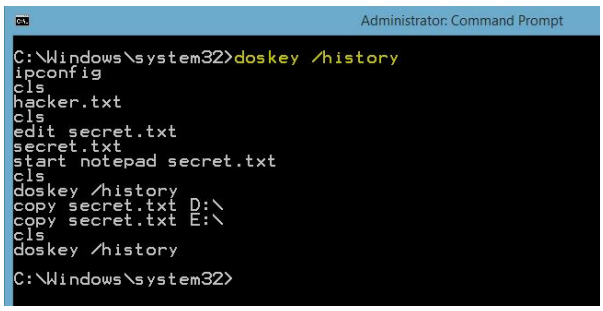

Figure 1: Doskey utility in Windows command prompt.

To get the current uptime and system events and statistics of the local or remote system we can use a utility called Uptime2.exe. See Figure 2.

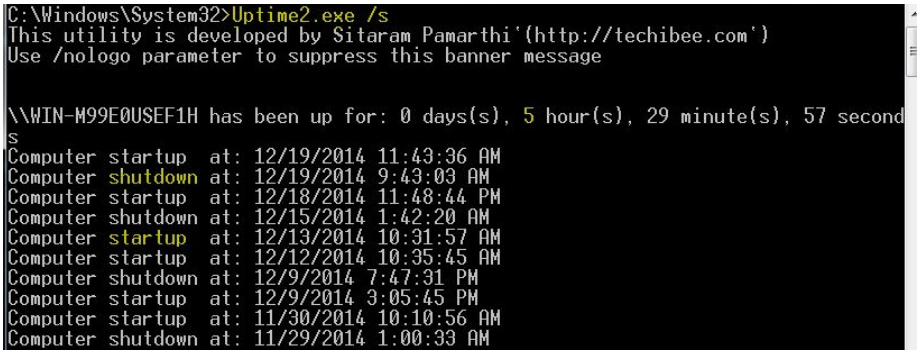
C:\Windows\System32>Uptime2.exe /s
This utility is developed by Sitaram Pamarthi'(http://techibee.com')
Use /nologo parameter to suppress this banner message
\\WIN-M99E0USEF1H has been up for: 0 days(s), 5 hour(s), 29 minute(s), 57 second
s
Computer startup at: 12/19/2014 11:43:36 AM
Computer shutdown at: 12/19/2014 9:43:03 AM
Computer startup at: 12/18/2014 11:48:44 PM
Computer shutdown at: 12/15/2014 1:42:20 AM
Computer startup at: 12/13/2014 10:31:57 AM
Computer startup at: 12/12/2014 10:35:45 AM
Computer shutdown at: 12/9/2014 7:47:31 PM
Computer startup at: 12/9/2014 3:05:45 PM
Computer startup at: 11/30/2014 10:10:56 AM
Computer shutdown at: 11/29/2014 1:00:33 AM

Figure 2: Uptime2.exe output giving uptimes for the windows system.

During an investigation we will always need to know who all were logged on to the system. Logging to a system can be remotely or locally. Information like these can add logical view to a context or a situation. The logs can be related to an event occurrence. Many tools are available like PsLoggedon, Netsessions, logonsessions etc. to learn the instantaneous information of the users. These tools can be downloaded from the *windows sysinternals* site. *Ps* tools in sysinternals are handy in many ways as such. See figure 3,4,5.

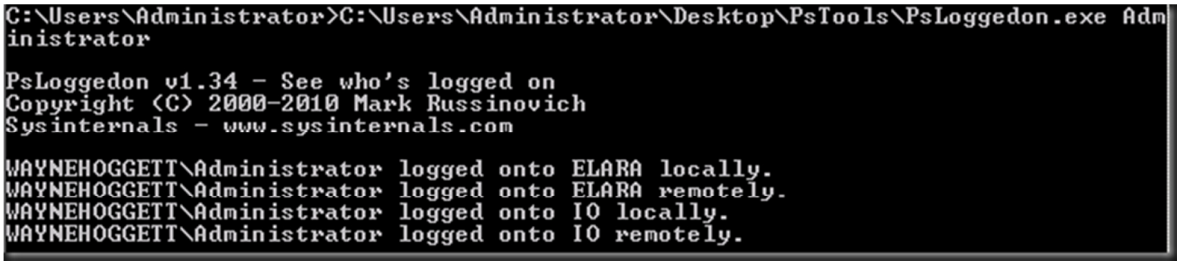
C:\Users\Administrator>C:\Users\Administrator\Desktop\PsTools\PsLoggedon.exe Adm
inistrator
PsLoggedon v1.34 - See who's logged on
Copyright (C) 2000-2010 Mark Russinovich
Sysinternals - www.sysinternals.com
WAYNEHOGGETT\Administrator logged onto ELARA locally.
WAYNEHOGGETT\Administrator logged onto ELARA remotely.
WAYNEHOGGETT\Administrator logged onto IO locally.
WAYNEHOGGETT\Administrator logged onto IO remotely.

Figure 3: PsLoggedon output.

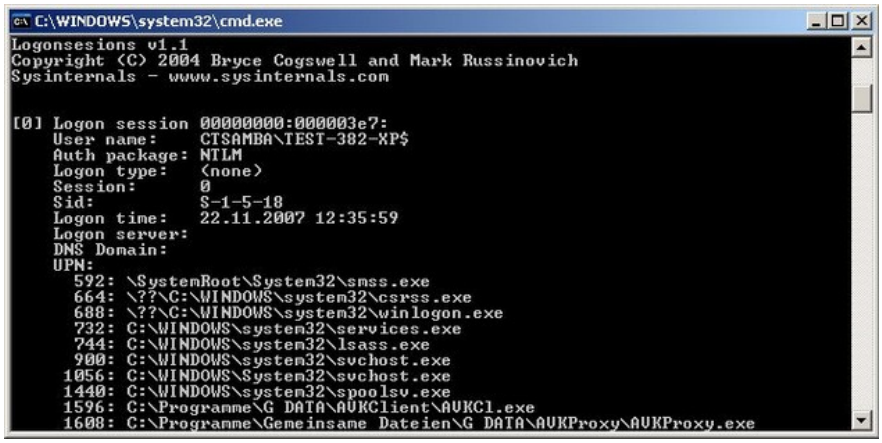

Figure 4:LoggedonSessions output.

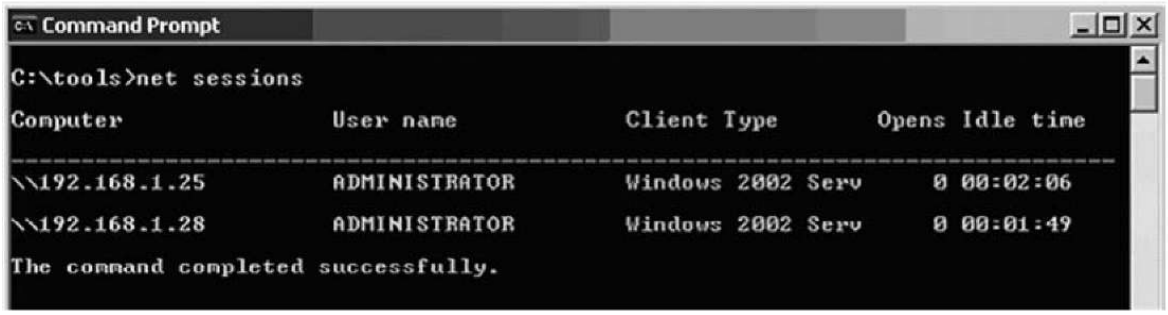
Command Prompt
C:\tools>net sessions
Computer User name Client Type Opens Idle time
\\192.168.1.25 ADMINISTRATOR Windows 2002 Serv 0 00:02:06
\\192.168.1.28 ADMINISTRATOR Windows 2002 Serv 0 00:01:49
The command completed successfully.

Figure 5:Net Sessions output.

Similarly we can also get which files were open at the time of logged users. This is also important many times as to co-relate whether which users were probably using which files of the system. Tools that can be used to access information pertaining opened files are: Netfile, PsFile, open files etc. Figure 6 gives a taste of how these tools can give out information of files that we open in the system.

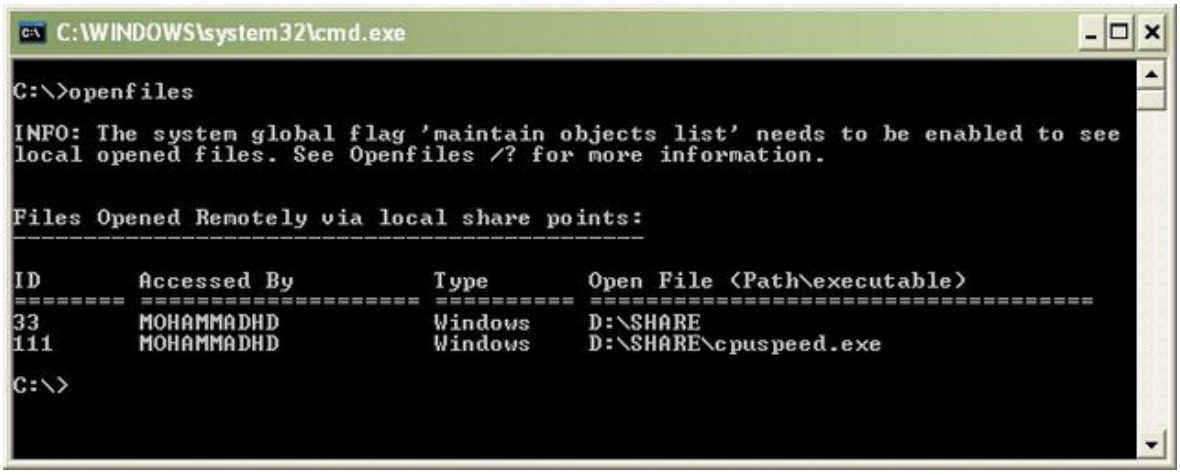
C:\WINDOWS\system32\cmd.exe
C:\>openfiles
INFO: The system global flag 'maintain objects list' needs to be enabled to see local opened files. See Openfiles /? for more information.
Files Opened Remotely via local share points:
ID Accessed By Type Open File (Path\executable)
33 MOHAMMADHD Windows D:\SHARE
111 MOHAMMADHD Windows D:\SHARE\cpuspeed.exe
C:\>

Figure 6: openfiles output.

Tools like NetStat gives access to information partitioning current network connections to the host computer. This information will be lost over time and very difficult to trace as time passes by. Figure 7 gives an output of the NetStat command. Also, an investigator needs to discover what processes are running on the system. This system which can keep clues to a major crime in form of files or processes that are still on the acquired system is potentially used just before a crime. Information about processes like executable file path, commands to launch the process, time stamps, current modules etc. along with contexts needs to be collected. Tools like Tlist, Tasklist, Pslist, ListDlls etc. helps us to get all these information. Windows task manager does give some information but most of the time it does not show vital information, hence using above tools play significant role in forensics.

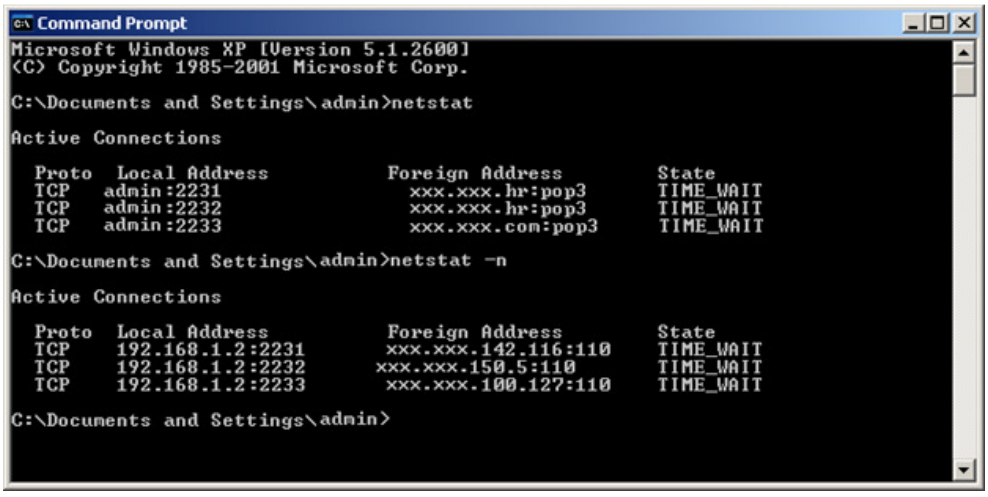

Figure 7: NetStat output.

Information about the status of the network interface cards (NIC) connected to a system can be very important. Wireless interfaces are very prominent these days and physical connection does not have too much presence. Hence, it's important to know the status of all interface devices (Network) is important. Tools like ipConfig, promiscDetect, promgry helps in getting the vital information (see Figure 8, Figure 9, Figure 10). Clipboards of windows are another aspect which is of utmost importance to the investigators. Clipboards contain latest copied area of memory which can be for later use. Clipboards facilitate users to move data in some way between documents or applications. The fact that recently copied and pasted items do remain on clipboard can give clue to vital evidences or circumstances leading to a crime. pclip is a command-line utility which helps the investigators to retrieve contents of a clipboard.

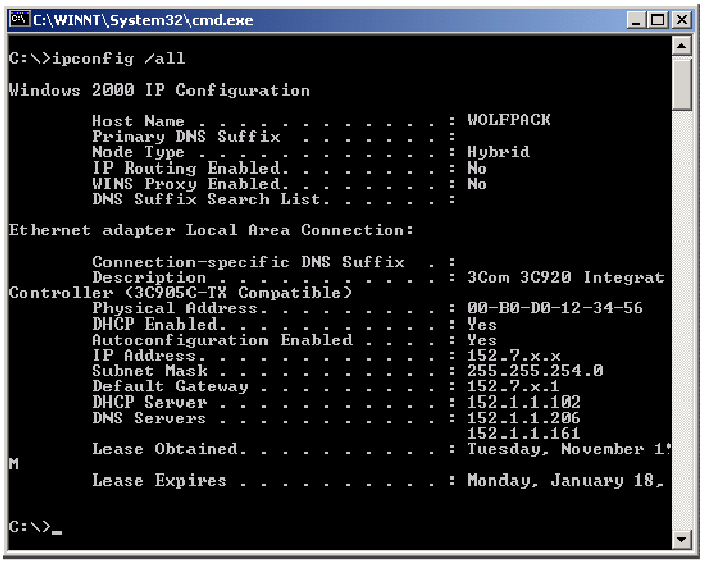
C:\WINNT\System32\cmd.exe
C:\>ipconfig /all
Windows 2000 IP Configuration
Host Name . . . . . : WOLFPACK
Primary DNS Suffix . . . . . :
Node Type . . . . . : Hybrid
IP Routing Enabled. . . . . : No
WINS Proxy Enabled. . . . . : No
DNS Suffix Search List. . . . . :
Ethernet adapter Local Area Connection:
Connection-specific DNS Suffix . :
Description . . . . . : 3Com 3C920 Integrat
Controller (3C905C-TX Compatible)
Physical Address. . . . . : 00-B0-D0-12-34-56
DHCP Enabled. . . . . : Yes
Autoconfiguration Enabled . . . . : Yes
IP Address. . . . . : 152.7.x.x
Subnet Mask . . . . . : 255.255.254.0
Default Gateway . . . . . : 152.7.x.1
DHCP Server . . . . . : 152.1.1.102
DNS Servers . . . . . : 152.1.1.206
152.1.1.161
Lease Obtained. . . . . : Tuesday, November 1
M
Lease Expires . . . . . : Monday, January 18,
C:\>

Figure 8: one of the output of ipConfig command.

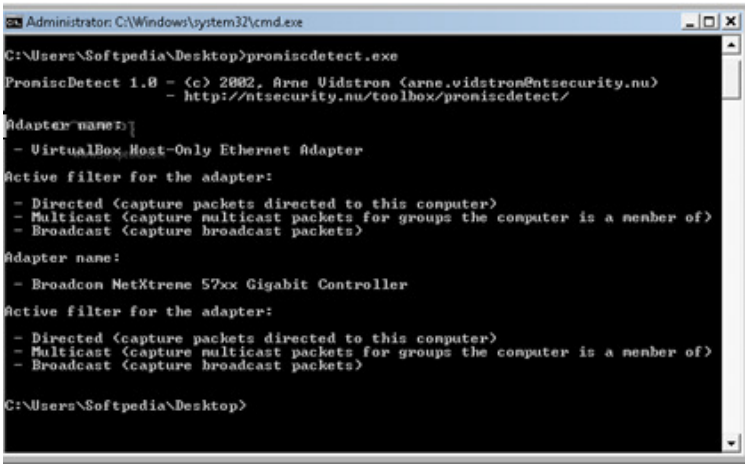
Administrator: C:\Windows\system32\cmd.exe
C:\Users\Softpedia\Desktop>promiscdetect.exe
PromiscDetect 1.0 - (c) 2002, Arne Vidstrom (arne.vidstrom@ntsecurity.nu)
- http://ntsecurity.nu/toolbox/promiscdetect/
Adapter name:
- VirtualBox Host-Only Ethernet Adapter
Active filter for the adapter:
- Directed (capture packets directed to this computer)
- Multicast (capture multicast packets for groups the computer is a member of)
- Broadcast (capture broadcast packets)
Adapter name:
- Broadcom NetXtreme 57xx Gigabit Controller
Active filter for the adapter:
- Directed (capture packets directed to this computer)
- Multicast (capture multicast packets for groups the computer is a member of)
- Broadcast (capture broadcast packets)
C:\Users\Softpedia\Desktop>

Figure 9:Promiscdetect command.

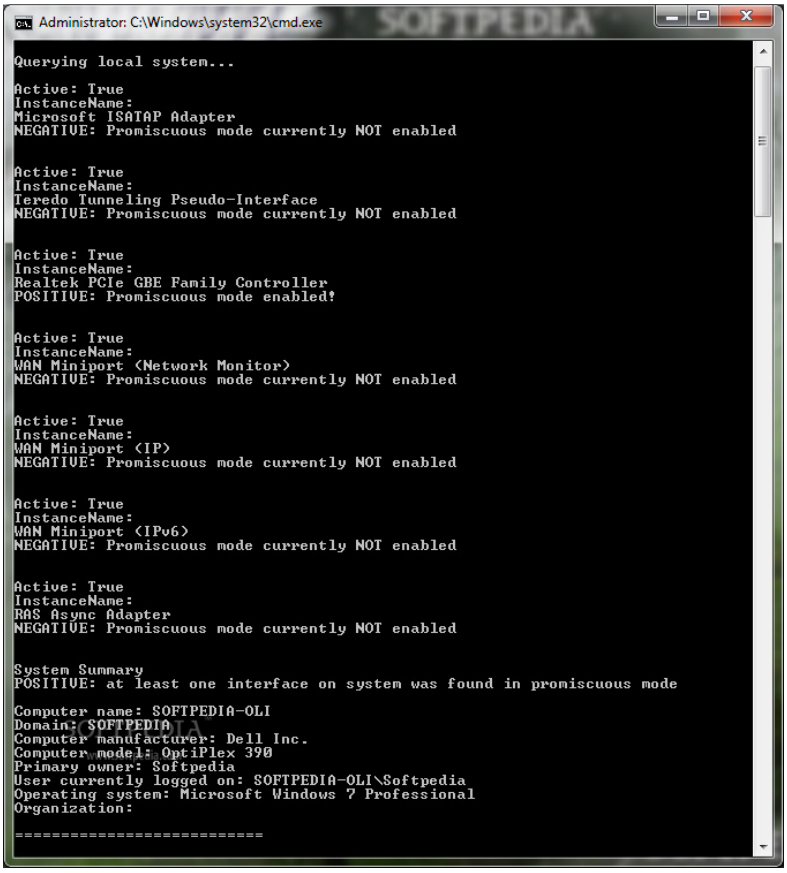

Figure 10:Promqry Command output.

C:\promqry>promqry /?
Promqry version 1.0 usage
Queries system(s) for interfaces
running in promiscuous mode
To query local system's interfaces run:
promqry.exe
notes: returns zero if any interfaces found in promiscuous mode
returns 1 if no interfaces found in promiscuous mode
returns 99 if error encountered
-np and -nv options are not valid for local query
To query a remote system's interfaces run:
promqry.exe remote_IP | remote_name [-nv]
notes: returns zero if any interfaces found in promiscuous mode
returns 1 if no interfaces found in promiscuous mode
returns 99 if error encountered
-nv means no verbose output, only reports errors and
systems with interfaces in promiscuous mode
To query a range of remote systems' interfaces run:
promqry.exe start_remote_IP:end_remote_IP [-np] [-nv]
notes: start_remote_IP must be lower than end_remote_IP
-np means no ping before query
-np only valid when querying range of systems
-nv means no verbose output, only reports errors and
systems with interfaces in promiscuous mode

Figure 11: various options with promqry.

Several other information like; mapped drives, shares or stored folders etc. also needs to be collected for future tests and analysis. Mapped drives to a system are those which the user has created. These information are volatile but can be correlated to network connections or drive activities leading to a crime. A system resources can be shared in many different ways like shared folders, shared network access etc. This information can be retrieved in many ways like scanning the registry for shares. Also, command like 'share' can be used for the same.

1.2.2.2 Non Volatile information

Non-volatile information remains on a secondary storage device and persists even after power is off. This information can be collected later on after all perishable information (volatile) can be collected after the seizure of the system. Investigators can collect these information after procuring the device and doing all the formalities of the seizure/procuring/capturing the device under law so that the discoveries later on does not get laid down during hearing. Using command line 'dir /o: d' the examiner can list out the recent updates that is listed by the command.

Registry information

Registry information can have a good impact on the forensic analysis and investigation. Tools like *reg* (see figure12,13) and *regedit* (see figure 14) helps in to get registry entries via important keys. Few important keys important keys present in registries are runMRU, startup objects, last accessed key, addresses in internet explorer, last saved directory in internet explorer.

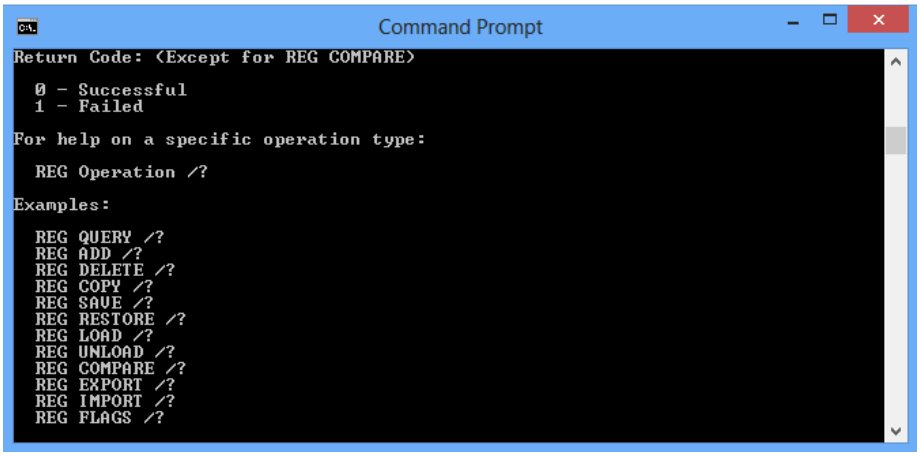

Figure 12:Options in reg tool.

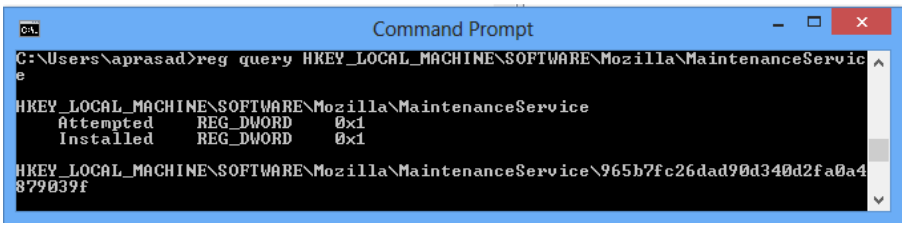

Figure 13: example output of reg.

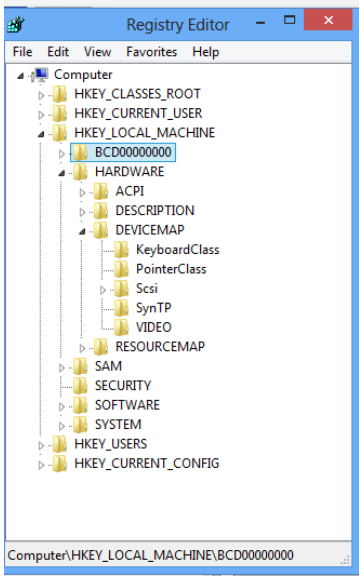

Figure 14:regedit command in windows.

RunMRU stores information about recently typed commands from run window, startup objects are those objects or apps that start automatically on startup in windows.

Key for RunMRU is:
Computer\HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Explorer\RunMRU

Key for startup object is:
Computer\HKCU\<SID>\SOFTWARE\Microsoft\Windows\CurrentVersion\Run
Computer\HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\Run
To access the least accessed key in registry use key:
Computer\HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Applets\Regedit
To get last typed urls in internet explorer use key:
Computer\HKEY_CURRENT_USER\Software\Microsoft\Internet Explorer\TypedUrls
To get last saved directory in IE use key:
Computer\HKEY_CURRENT_USER\Software\Microsoft\Internet Explorer Download Directory
To get security ids Microsoft use:
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\WindowsNT\CurrentVersion\ProfileList key
Another area of registry which has valuable information for forensics analysis is the protected storage area. These storages are encrypted. However, we can get access to these areas using tools like *Access Data tool kit* (see figure 15).

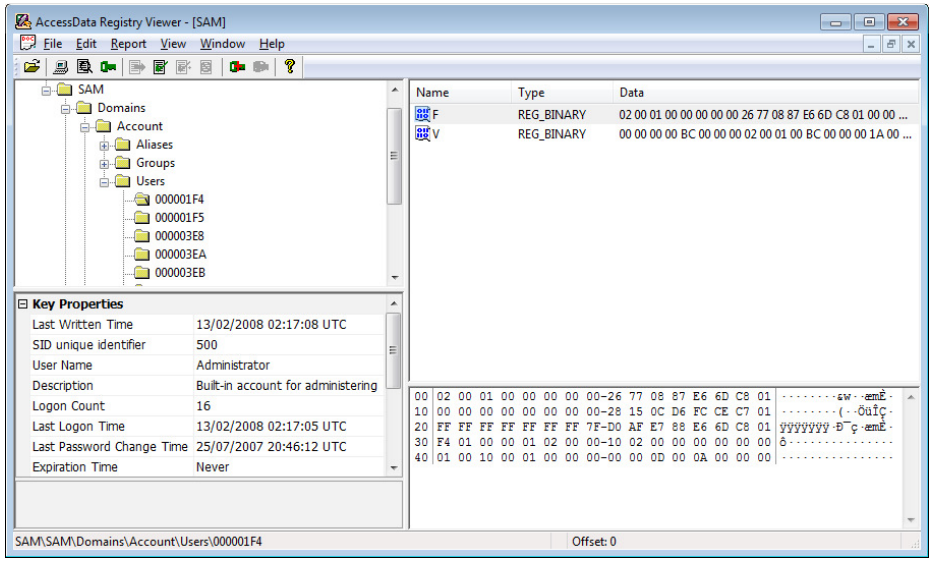
AccessData Registry Viewer - [SAM]
File Edit Report View Window Help
SAM
Domains
Account
Aliases
Groups
Users
000001F4
000001F5
000003E8
000003EA
000003EB
Name Type Data
F REG_BINARY 02 00 01 00 00 00 00 00 26 77 08 87 E6 6D C8 01 00 00 ...
V REG_BINARY 00 00 00 00 BC 00 00 00 02 00 01 00 BC 00 00 00 1A 00 ...
Key Properties
Last Written Time 13/02/2008 02:17:08 UTC
SID unique identifier 500
User Name Administrator
Description Built-in account for administering
Logon Count 16
Last Logon Time 13/02/2008 02:17:05 UTC
Last Password Change Time 25/07/2007 20:46:12 UTC
Expiration Time Never
SAM\SAM\Domains\Account\Users\000001F4
Offset: 0

Figure 15:AccessData tool window.

Some time it may be very important to get record what are all the devices that were connected to a system. To gain access to this information we can use tools like (windows device console) *DevCon* of Microsoft. Device manager of windows is also available for some Figure 16 shows the output of *DevCon*.

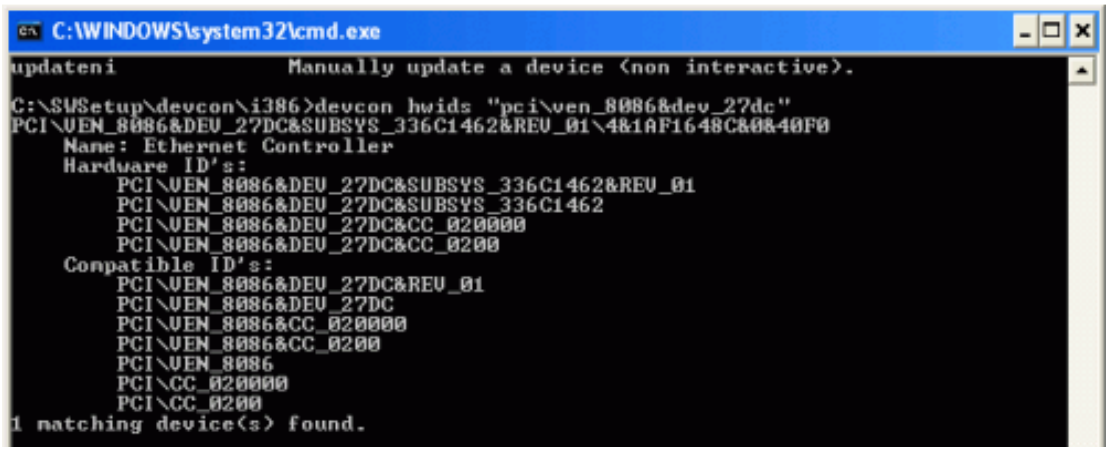

Figure 16: Devcon output.

1.3 RECOVERING DELETED FILES AND PARTITIONS

1.3.1 Anatomy of a disc drive

A hard disk drive (HDD), hard disk, hard drive or fixed disk is a data storage device used for storing and retrieving digital information using one or more rigid ("hard") rapidly rotating disks (platters) coated with magnetic material. The platters are paired with magnetic heads arranged on a moving actuator arm, which read and write data to the platter surfaces. Data is accessed in a random-access manner, meaning that individual blocks of data can be stored or retrieved in any order rather than sequentially. HDDs retain stored data even when powered off. The primary characteristics of an HDD are its capacity and performance. Capacity is specified in unit prefixes corresponding to powers of 1000: a 1-terabyte (TB) drive has a capacity of 1,000 gigabytes (GB; where 1 gigabyte = 1 billion bytes). Typically, some of an HDD's capacity is unavailable to the user because it is used by the file system and the computer operating system, and possibly inbuilt redundancy for error correction and recovery.

An HDD records data by magnetizing a thin film of ferromagnetic material on a disk. Sequential changes in the direction of magnetization represent binary data bits. The data is read from the disk by detecting the transitions in magnetization. User data is encoded using an encoding scheme, such as run-length limited encoding, which determines how the data is represented by the magnetic transitions.

In computer disk storage, a *sector* is a subdivision of a track on a magnetic disk or optical disc. Each sector stores a fixed amount of user-accessible data, traditionally 512 bytes for hard disk drives (HDDs) and 2048 bytes for CD-ROMs and DVD-ROMs. Newer HDDs use 4096-byte (4 KB) sectors, which are known as the Advanced Format (AF).

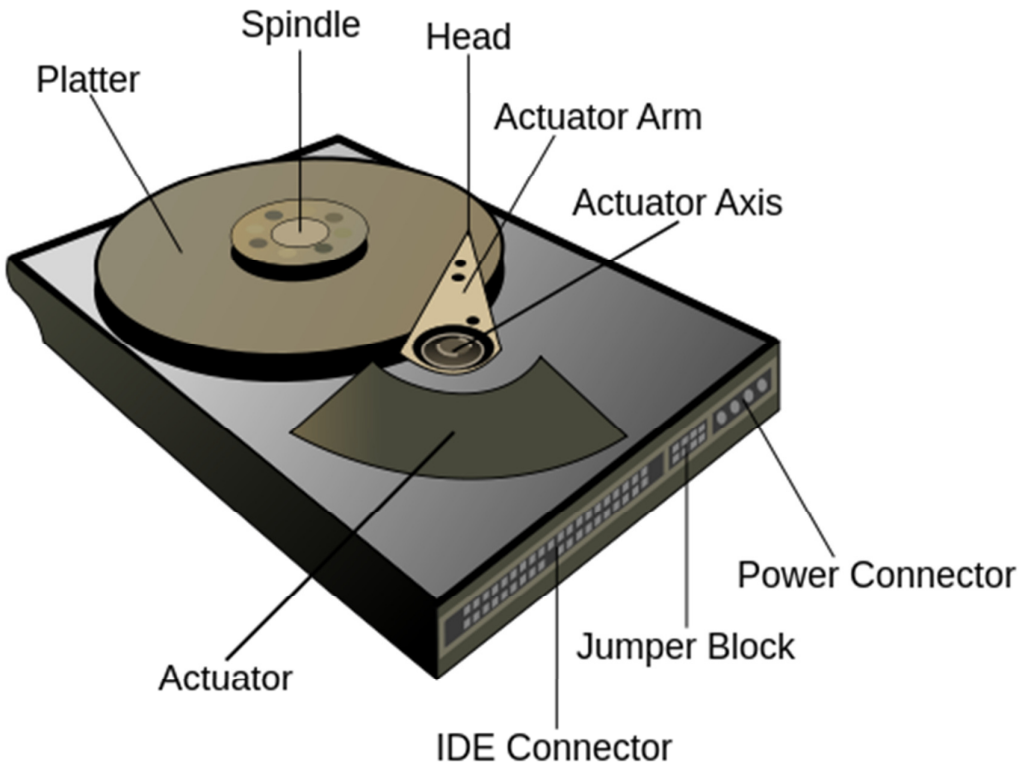

Figure 17: Hard Disk parts

Geometrically, the word sector means a portion of a disk between a center, two radii and a corresponding arc (see Figure 17, item B), which is shaped like a slice of a pie. Thus, the *disk sector* (Figure 17, item C) refers to the intersection of a *track* and geometrical *sector*.

In disk drives, each physical sector is made up of three basic parts, the sector header, the data area and the error-correcting code (ECC). The sector header contains information used by the drive and controller; this information includes sync bytes, *address identification*, flaw flag and header parity bytes. The header may also include an alternate address to be used if the data area is undependable. The *address identification* is used to ensure that the mechanics of the drive have positioned the read/write head over the correct location. The data area contains the recorded user data, while the ECC field contains codes based on the data field, which are used to check and possibly correct errors that may have been introduced into the data.

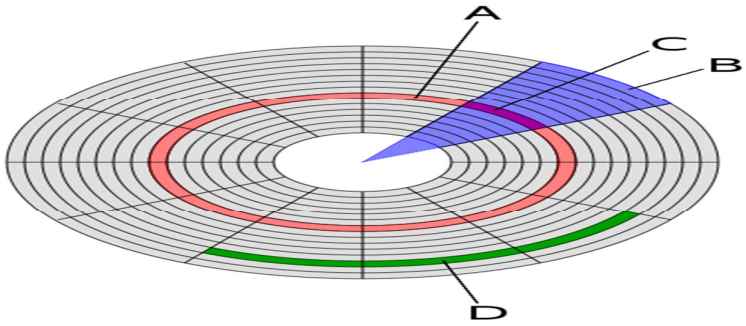

Figure 18: disk layout

1.3.2 Data organization in Windows

Windows organisation data using following structures or elements:

Cluster: Group of sectors form a cluster. Typically clusters can be of 32 kB. Clusters use logical representation of sectors.

Partition: Logical division of the physical storage. A large physical storage needs to be partitioned in smaller size so that the OS can use each partition as separate entity. However, smart user hides data into hidden or temporarily deleted partitions.

Master Boot Record (MBR): Every OS starts with reading a boot record or programme at the first location of a partition that is mapped to the OS hardware but up sequence.

FAT32: Initially FAT was the widely used allocation systems. FAT stands for file allocation table and it's a structure that keeps vital Meta data of a file that resides on the hard disc or any other storage. The FAT system uses a certain defines mechanisms to construct/store a file. These pre-defined mechanisms are used very nicely by forensics tools to reconstruct file by forensic tools.

New Technology File System (NTFS): The NTFS is a latest standard version introduced by Microsoft which is little advanced in terms of the allocation table structure as well as keeps the data compatible foe other OS to work upon. NTFS is currently used with Window OS. A file in NTFS is deleted in two steps.

a. The file moved to recycle bin and the meta-data is updated. These meta-data can be read by the forensic tools.
b. The clusters occupied (originally) by the file are then marked available for new file and the master file table is also updated. When the user empties the recycle bin the NTFS marks the clusters as available and considers the files to be permanently deleted.

1.3.3 Retrieving deleted files

When a file is deleted, the file system removes the file logically. That is, it removes all the meta-data and stamps related to the file. However, the file still resides in the disk as a physical entity until it is overwritten. These physical areas can be very easily explored and read and converted to a readable file using forensic application. It is observed that data resides on a computer for a very long time and are retrieved to a good extent.

1.3.4 Retrieving cached files

One can find the webpage visited by the suspect or the victim by looking into the cache. The cache file of an application can be spread across in the system storage. We can confine only search by using typical keywords elated to the case or probable websites.

1.3.5 Retrieving files in unallocated space

In general a deleted file can be searched sequentially or structurally by looking for file headers or extensions. However, certain tools help us to scan and look for broken headers and use supplementary headers to retrieve data or at least retrieve blocks of a lost file for

unallocated space. These retrieved blocks can later be studied and reformed using other tools to retrieve lost files to a great extent. This is also called as file carving.

Meta data of the files can be found from the applications used to create the files however there can be certain tools available to view the metadata of a files like Meta Viewer, Metadata Analysis, iscrub etc.

1.4 MORE ABOUT RECOVERING LOST FILES/DATA

1.4.1 Slack space, swap file, deleted files

Even these days most of the users aren't careful and thus the forensic experts get more clues because of this. The user's ignorance of how computers manage memory, disks and related stuff leaves lots of spaces which are rather invisible to the user (who can be a subject of an investigation). Let us look at three potential locations where an investigator explores to find lost data as deleted files and slack space, swap space etc.

1.4.1.1 Slack Space

Slack Space is the unused space in a disk cluster. The DOS and Windows file systems use fixed-size clusters. Even if the actual data being stored requires less storage than the cluster size, an entire cluster is reserved for the file. The unused space is called the slack space.

DOS and older Windows systems use a 16-bit file allocation table (FAT), which results in very large cluster sizes for large partitions. For example, if the partition size is 2 GB, each cluster will be 32 K. Even if a file requires only 4 K, the entire 32 K will be allocated, resulting in 28 K of slack space. In computer forensics, slack space is examined because it may contain meaningful data.

1.4.1.2 Swap space

Swap space is the area on a hard disk which is part of the Virtual Memory of your machine, which is a combination of accessible physical memory (RAM) and the swap space. Swap space temporarily holds memory pages that are inactive. Swap space is used when your system decides that it needs physical memory for active processes and there is insufficient unused physical memory available. If the system happens to need more memory resources or space, inactive pages in physical memory are then moved to the swap space therefore freeing up that physical memory for other uses. On a Windows machine, the swap space is a file called Pagefile.sys.

Almost everything on a RAM can be swapped if necessary, because of this we can find very important and forensically interesting things in the swap space. Apart from plain-text data of an encrypted text in a disk file we can even find encryption keys! Thanks to flaw-full weaknesses in some applications that allow unencrypted keys to reside in memory. Also, part of e-mails or matter stored at remote locations might still reside in swap space. And to relief of all investigators, any standard disk maintenance utility can access this information.

On Windows, the swap file is a hidden file found in the root directory called pagefile.sys. The registry path for the swap file is (can be subject to change):

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\SessionManager\Memory Management.

Several tools are available to retrieve slack space and swap space on windows system. Slack space can retrieved using a popular tools like DriveSpy, Encase etc. The overall process of retrieving involves following steps:

a) Connect to the computer.
b) Have a bit level image of original media.
c) Keep a hashed value of all images.
d) Use key word searches and hash analysis etc. using tools like Encase.

Tools like DriveSpy can be used to do some of the above processes.

1.4.1.3 File Carving

File carving can be used to recover data from a hard disk where the metadata is missing or damaged, especially by professional data recovery companies.

When a file is deleted, only the entry in the file system metadata is removed, while the actual data is still on the disk. After a format and even a repartitioning it might be that most of raw data is untouched and can be recovered using file carving.

All file systems contain some metadata that describes the actual file system. At a minimum the following is stored: the hierarchy of folders and files, with names for each. For each file is also stored the physical address on the hard disk where the file is stored. As explained below, a file might be scattered in fragments at different physical addresses.

File carving is the process of trying to recover files without this metadata. This is done by analyzing the raw data and identifying what it is (text, executable, png, mp3, etc.). This can be done in different ways, but the simplest is to look for headers. For instance, every Java class file has as its first four bytes the hexadecimal value CA FE BA BE. Some files contain footers as well, making it just as simple to identify the ending of the file.

Most file systems, such as FAT and UNIX Fast File System, work with the concept of clusters of an equal and fixed size. For example, a FAT32 file system might be broken into clusters of 4 KB each. Any file smaller than 4 KB fits into a single cluster, and there is never more than one file in each cluster. Files that take up more than 4 KB are allocated across many clusters. Sometimes these clusters are all contiguous, while other times they are scattered across two or potentially many more so called fragments, with each fragment containing a number of contiguous clusters storing one part of the file's data. Obviously large files are more likely to be fragmented.

File carving is a highly complex task, with a potentially huge number of permutations to try. To make this task tractable, carving software typically makes extensive use of models and heuristics. This is necessary not only from a standpoint of execution time, but also for the accuracy of the results. State of the art file carving algorithms use statistical techniques like sequential hypothesis testing for determining fragmentation points.

1.4.1.4 Event logs

Event logs are stored in Metadata files. The entries in these files can be retrieved on a good way depending upon how refining is carried out by investigators. The victim or suspect system log entries change rapidly as the new events are recorded. The event logs can also be configured minimal to maximum events and durations. We can use tools like Ps log list and EVT to retrieve event records. See figure 19,20, 21.

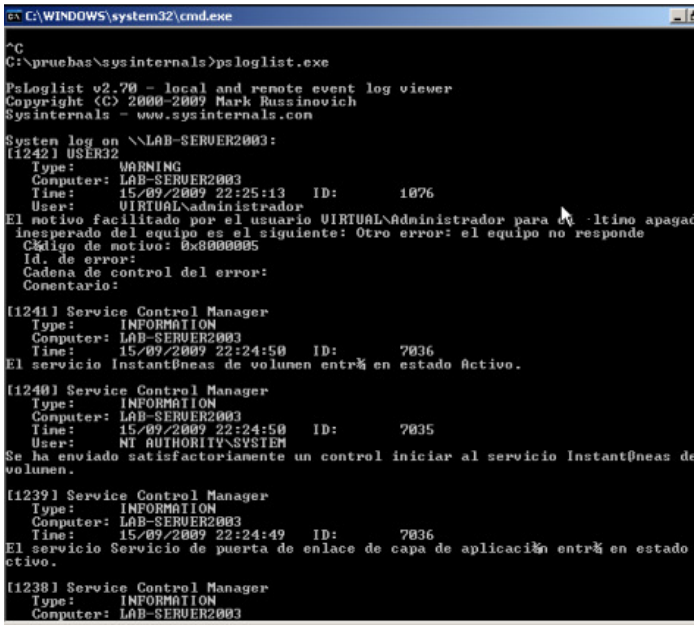
C:\WINDOWS\system32\cmd.exe
^C
C:\pruebas\sysinternals>psloglist.exe
PsLoglist v2.70 - local and remote event log viewer
Copyright (C) 2000-2009 Mark Russinovich
Sysinternals - www.sysinternals.com
System log on \\LAB-SERVER2003:
[1242] USER32
Type: WARNING
Computer: LAB-SERVER2003
Time: 15/09/2009 22:25:13 ID: 1076
User: VIRTUAL\administrador
El motivo facilitado por el usuario VIRTUAL\Administrador para el último apagado
inesperado del equipo es el siguiente: Otro error: el equipo no responde
Código de motivo: 0x8000005
Id. de error:
Cadena de control del error:
Comentario:
[1241] Service Control Manager
Type: INFORMATION
Computer: LAB-SERVER2003
Time: 15/09/2009 22:24:50 ID: 7036
El servicio Instant@neas de volumen entr% en estado Activo.
[1240] Service Control Manager
Type: INFORMATION
Computer: LAB-SERVER2003
Time: 15/09/2009 22:24:50 ID: 7035
User: NT AUTHORITY\SYSTEM
Se ha enviado satisfactoriamente un control iniciar al servicio Instant@neas de
volumen.
[1239] Service Control Manager
Type: INFORMATION
Computer: LAB-SERVER2003
Time: 15/09/2009 22:24:49 ID: 7036
El servicio Servicio de puerta de enlace de capa de aplicaci%n entr% en estado
ctivo.
[1238] Service Control Manager
Type: INFORMATION
Computer: LAB-SERVER2003

Figure 19: PsLoglist output.

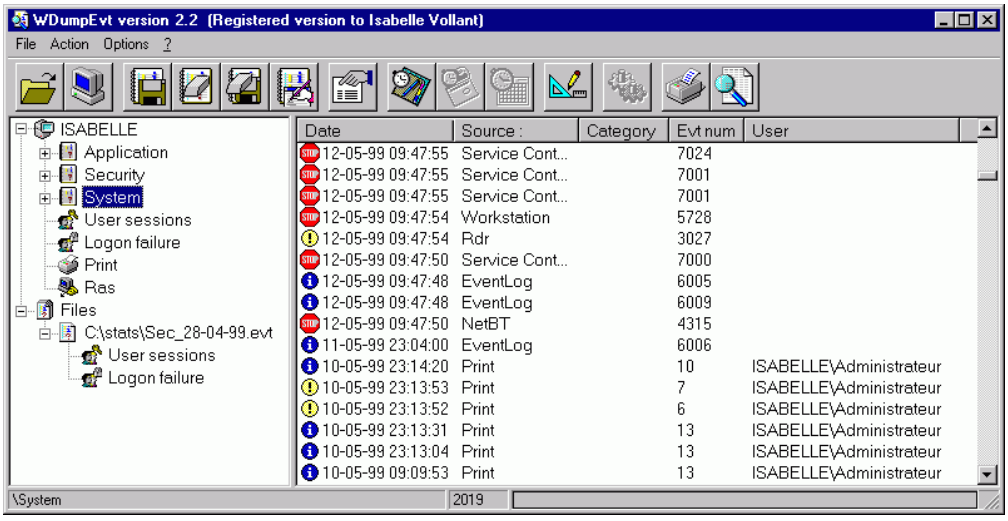

Figure 20: WDumEvt window (showing system).

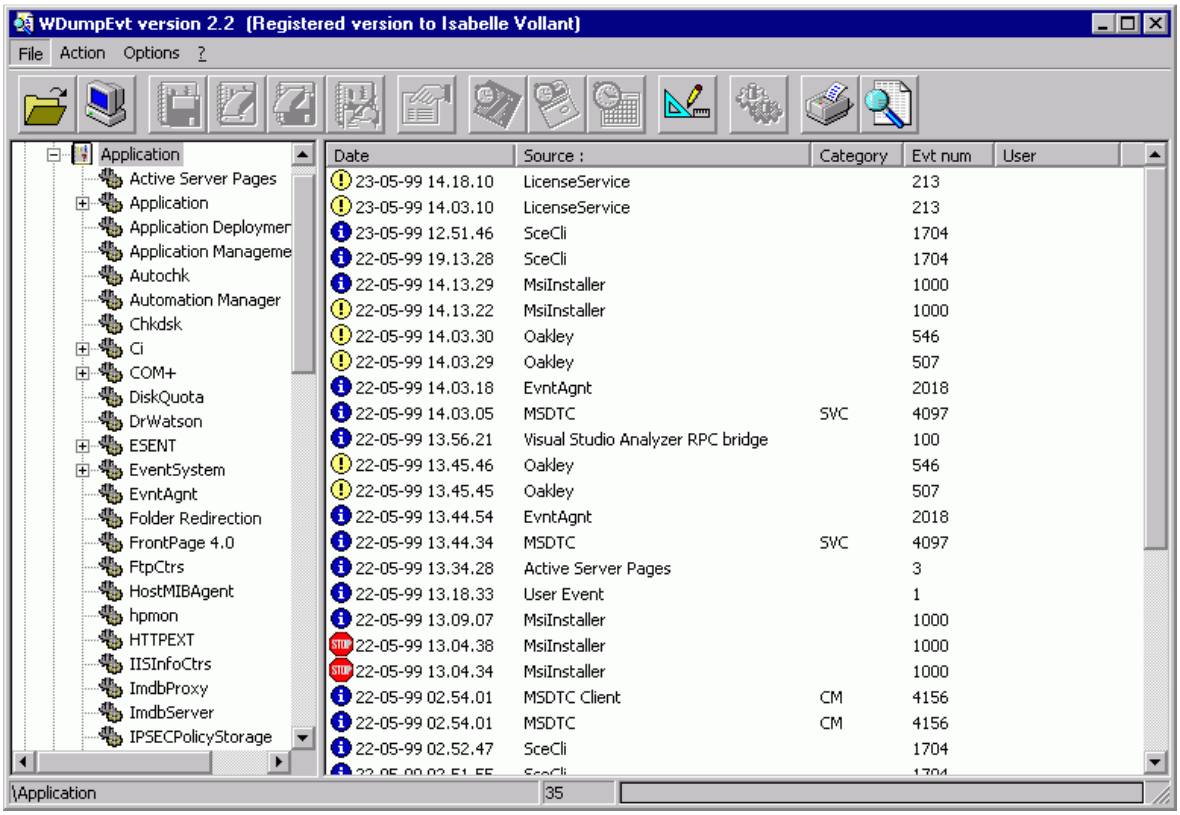

Figure 21: WDumEvt window (showing apps).

1.5 SUMMARY

1. Digital Forensic examiners will most likely encounter Windows and will have to collect evidence from it in almost all cyber-crime cases.
2. Forensics can be looked as Volatile and non-volatile information in Windows.
3. Volatile Information in Windows can disappear or be easily modified.
4. Non-volatile information remains on a secondary storage device and persists even after power is off.
5. Major tools for extracting volatile information are PsLoggedon, Netsessions, logonsessions, doskey uptime etc.
6. Tools that can be used to access information pertaining opened files are: Netfile, PsFile, open files etc.
7. Tools like NetStat gives access to information partitioning current network connections to the host computer
8. Tools like ipConfig, promiscDetect, promgry helps in getting the vital information pertaining network interface cards (NIC) in Windows.
9. Registry information can have a good impact on the forensic analysis and investigation and collecting these information would be very vital.
10. Tools like *DevCon, Access Data tool kit, reg* and *regedit* helps in extracting non-volatile information in Windows.
11. Windows organises data using structures or elements like *Cluster, Partition, Master Boot Record, FAT32, New Technology File System.*

12. Files that are deleted, lost, cached or unallocated can be retrieved using various methods and tools.

1.6 CHECK YOUR PROGRESS

1. Fill in the blanks.

a) The Microsoft windows versions that are currently in use are _____ and _____.
b) Major forensics areas in windows are _____ and _____ information
c) _____is a utility for DOS and Microsoft Windows that adds command history
d) _____ in windows contain latest copied area of memory which can be for later use.
e) Tools like *reg* and *regedit* helps in to get _____ via important keys.
f) In computer disk storage, a _____is a subdivision of a track on a magnetic disk or optical disc.
g) _____ is the amount of on-disk file space from the end of the logical record information to the end of the physical disk record.
h) _____ is the process of trying to recover files without a file system metadata.

2. State True or False.

a) Volatile Information retains its contents even when the power is interrupted the stored data is immediately lost.
b) System time, logged users, open files, network information and drives that are mapped shared folders are examples of non-volatile information in windows.
c) Registry information is an example of volatile information
d) Group of sectors form a cluster.
e) When a file is deleted, the file system removes the file logically i.e. it removes all the meta-data and stamps related to the file.

1.7 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) Windows 8 and Windows 10.
b) Volatile and non-volatile.
c) Doskey.
d) Clipboards.
e) Registry entries.
f) Sector.
g) Slack space.
h) File carving.

2. State True or False

a) (F)
b) (F)
c) (F)

d) (T)
e) (T)

1.8 FURTHER READINGS

- Windows Forensic Analysis Toolkit, Third Edition: Advanced Analysis Techniques for Windows 73rd Edition, by Harlan Carvey.
- File system forensic analysis 1st edition, by Brian carrier
- http://www.gfi.com/blog/top-20-free-digital-forensic-investigation-tools-for-sysadmins/
- Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
- Investigating Hard Disks, File and Operating Systems: EC-Council | Press

1.9 MODEL QUESTIONS

1. State the usage and forensic importance of PsLoggedon, Netsessions, logonsessions tools.
2. How the deleted and lost files are recovered in a windows system?
3. Describe the disk and file structure in a windows system.
4. What is a slack space, swap space and file carving?
5. How is registry information important in windows forensics?

References, Article Source & Contributors

[1] Disk Sector, https://en.wikipedia.org/wiki/Disk_sector, retrieved Nov 2015
[2] DriveSpy, https://www.digitalintelligence.com/software/disoftware/drivespy/, retrieved Nov 2015
[3] File Carving, https://en.wikipedia.org/wiki/File_carving, retrieved Nov 2015
[4] Hard Disk Drive, https://en.wikipedia.org/wiki/Hard_disk_drive, retrieved Nov 2015
[5] Operating Systems, https://en.wikipedia.org/wiki/Operating_system, retrieved Nov 2015
[6] What is slack space, A Webopedia Definition, www.webopedia.com/TERM/S/slack_space

Bibliography

[1] Windows System artefacts, http://resources.infosecinstitute.com/windows-systems-and-artifacts-in-digital-forensics-part-i-registry/, retrieved Nov 2015
[2] Tom Olzak, IT Security, http://www.techrepublic.com/blog/it-security/computer-forensics-finding-hidden-data/, May 21, 2007, retrieved Nov 2015.

UNIT II: LOGS & EVENT ANALYSIS AND PASSWORD CRACKING

2.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand registry and fetch various keys in registry related to event logs.
- Understand the event log file structure and know how event information can be retrieved from log files correlate its use while doing forensic investigation.
- Know user account policies, audit policies and mechanisms of changing audit policy and correlate its use while doing forensic investigation.
- Know various tools used for log and event analysis.
- Know basic ways passwords are stored in Windows and correlate while doing forensic investigation.
- Understand various ways of password attacks (password hacking) and correlate while doing forensic investigation.
- Know various tools for password hacking useful in forensic investigation.

2.2 INTRODUCTION

In this chapter we will discuss two very important aspects of windows and other systems which plays very vital role in forensics. They are: Event logs and Password cracking. In computer log management and intelligence, log analysis (or system and network log analysis) is an art and science seeking to make sense out of computer-generated records (also called log or audit trail records). The process of creating such records is called data logging. Typical reasons why people perform log analysis are:

- Compliance with security policies
- Compliance with audit or regulation
- System troubleshooting
- Forensics (during investigations or in response to subpoena)
- Security incident response

The Security Log, in Microsoft Windows, is a log that contains records of login/logout activity or other security-related events specified by the system's audit policy. Auditing allows administrators to configure Windows to record operating system activity in the Security Log. Event logging provides system administrators with information useful for diagnostics and auditing. The different classes of events that will be logged, as well as what details will appear in the event messages, are often considered early in the development cycle. Many event logging technologies allow or even require each class of event to be assigned a unique "code", which is used by the event logging software or a separate viewer (e.g., Event Viewer) to format and output a human-readable message. This facilitates localization and allows system administrators to more easily obtain information on problems that occur.

Windows registry is also a very important source to maintain and manage logs. As well registry also has variety of controls/keys where general records pertaining events etc. are maintained which can be very vital during digital forensics.

In cryptanalysis and computer security, password cracking is the process of recovering passwords from data that have been stored in or transmitted by a computer system. The purpose of password cracking might be to help a user recover a forgotten password (installing an entirely new password is less of a security risk, but it involves System Administration privileges), to gain unauthorized access to a system, or as a preventive measure by System Administrators to check for easily crack-able passwords. On a file-by-file basis, password cracking is utilized to gain access to digital evidence for which a judge has allowed access but the particular file's access is restricted.

2.3 WINDOWS REGISTRY

Windows registry keeps most of the information pertaining policies, status etc. in form of keys, sub keys and values. Windows registry can be worked upon by administrator through application like 'regedit'. Windows can also be supplied with a command like tool like 'reg' to help users work on registry. Registry contains hives under which sub keys are present. These hives play important role in the overall functioning of the system.

2.3.1 Registry and forensics

An investigator can acquire quite a good deal of information by studying and analysing registry. Many tools like ProDiscover, ProScript can be very handy to get a good deal of analysis of registry entries. Registry entries can be used to acquire and analyse many important information necessary for forensics analysis. These information use system, time zone, shares, audit policy, wireless SSIDS, auto start locations, user login, activities, USB removable devices, trusted devices, cache, cookie and history etc.

2.3.1.1 System information

Basic information of system can be acquired for registry. Certain system information and its registry key are listed below:

Table 3: Various important log attributes and respective registry keys.

System Information	Key
Computer Name	SYSTEM\CurrentControlSet\Control\ComputerName\ActiveComputerName
Time of last shutdown	SYSTEM\ControlSet00x\Control\Windows
Product name ,build, version etc.	SOFTWARE\Microsoft\Windows NT\CurrentVersion
Time zone settings	SYSTEM\CurrentControlSet\Control\TimeZoneInformation
User created shares	SYSTEM\CurrentControlSet\Services\lanmanserver\Shares
Audit policy	\SECURITY\Policy\PolAdtEv
Wireless SSIDs	SOFTWARE\Microsoft\WZCSVC\Parameters\Interfaces\{GUID}
USB devices connected	HKEY_LOCAL_MACHINE\System\CurrentControlSet\Enum\USBSTOR
last time	HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\DeviceClasses
Mounted Devices	HKEY_LOCAL_MACHINE\System\MountedDevices
User	SAM\SAM\Domains\Account\Users\{RID}

information stored in the user's NTUSER.DAT	Software\Microsoft\Windows\CurrentVersion\Explorer\UserAssist\{GUID}\Count
most recently used	\Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs
most recently used	\Software\Microsoft\Windows\CurrentVersion\Explorer\RunMRU
Search Assistant MRU Lists	Software\Microsoft\Search Assistant\ACMru
Internet downloads directory	Computer\HKEY_CURRENT_USER\Software\Microsof
Restore points	HKEY_LOCAL_MACHINE\Software\Microsoft\WindowsNT\CurrentVersion\SystemRestore

Table 3 list out few important keys and their paths. These information acquired using these keys has to be recorded using Encase and can lead to many conclusions while putting up the case.

'Computers' here is the name that the user gives to its computer. The name of computer generally is made once in the lifetime usage of the system and hence it can be used to trace various activities on network and internet carried by the user. Time of last shutdown is the time at which the system was completely shut down. This information can lead us to know the status of the user and time stamps of various files and can co-relate to give an idea of the mental status of the suspect. Sometime user themselves create shared folders and applications for others to use over local network or internet (remote desktops). This information can be traced out to find and analyse what kind of things or information the user was trying to share and thus stamps of the shared files/folders can also be analysed. Audit policy information can be very useful as it can let us know about what types of information/events an investigator should look for in the event log. Service set identifications (SSIDs) maintained by Windows can be useful in situations where unauthorised access is need to be investigated and IP addresses needs to be traced.

Artefacts of a USB devices connected to computer are also registered via PnP (plug and play) manager. The sub key formed for every USB device under the key path in table 1 is of the form *Disk &Ven_###&Prod_###&Rev###*. This and other information can be used to trace and collect vital evidences pertaining to a case. Similar is the case with mounted devices information under registry.

Many applications maintain MRU lists i.e. they keep a list of recently used files or opened/created files. Also search assistant MRU lists are also maintained by search applicants. MRU lists of connected systems etc. are also maintained. This information can of genuine help to understand victim's state of mind or condition just before the crime. System restore points can be studied to understand how and when the user created back-ups. Restore points can be used to understand long back status of the user work.

Events are any occurrences or triggering of an activity. The operating system logs some of these occurrences or events. However, the key PolAdEvt in registry can be used to set audit configuration in order to log events based on user requirements. Other key available for logging events is:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Eventlog\<Event Log>

One can view events logs from the control panel also (see *Figure 21*,*Figure 22* and *Figure 23*).

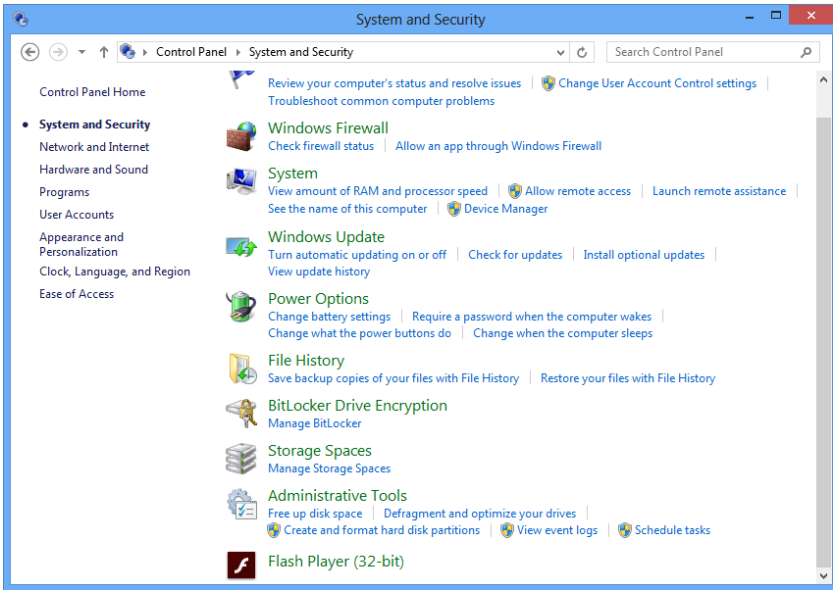

Figure 21: System and Security in control panel

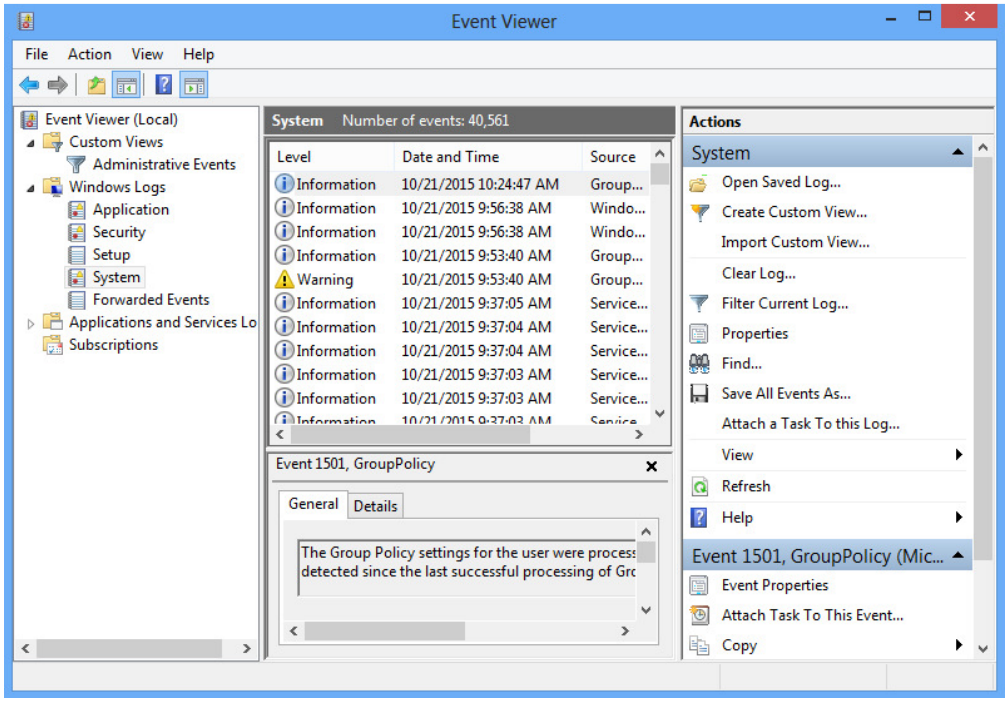

Figure 22: Event Viewer.

2.4 WINDOWS EVENT LOG FILE

In windows event logs are stored in binary format. Event logs are stored in form of headers and set of records. The event logs are in form of headers and set of records. The event logs are in form of pipe or buffer where event addition can lead to several of older events out of the file.

2.4.1 Windows Event Log File Format

Each log file consists of a Header record (given as ELF_LOGFILE_HEADER structure) and the Body. The body again consists of Event records, the Cursor record and unused space. The body could form a ring buffer, where the cursor record will mark the border between the oldest and the newest event record. Unused space could be empty, slack and padding

Windows Event Log (EVT)– ForensicsWiki,

www.forensicswiki.org/wiki/Windows_Event_Log_(EVT)

The Windows XML Event Log (EVTX) format was introduces in Windows Vista as a replacement for the Windows Event Log (EVT) format.

Whenever an event has to be written/created/updated ELF_LOGFILE_HEADER and the ELF_EOF_RECORD structures are written in the event log.

Whenever and application needs to log (or is set in registry to log an event) it calls ReportEvent function which adds an EVENTLOGRECORD structure taking the parameters from the system (see figure 3).

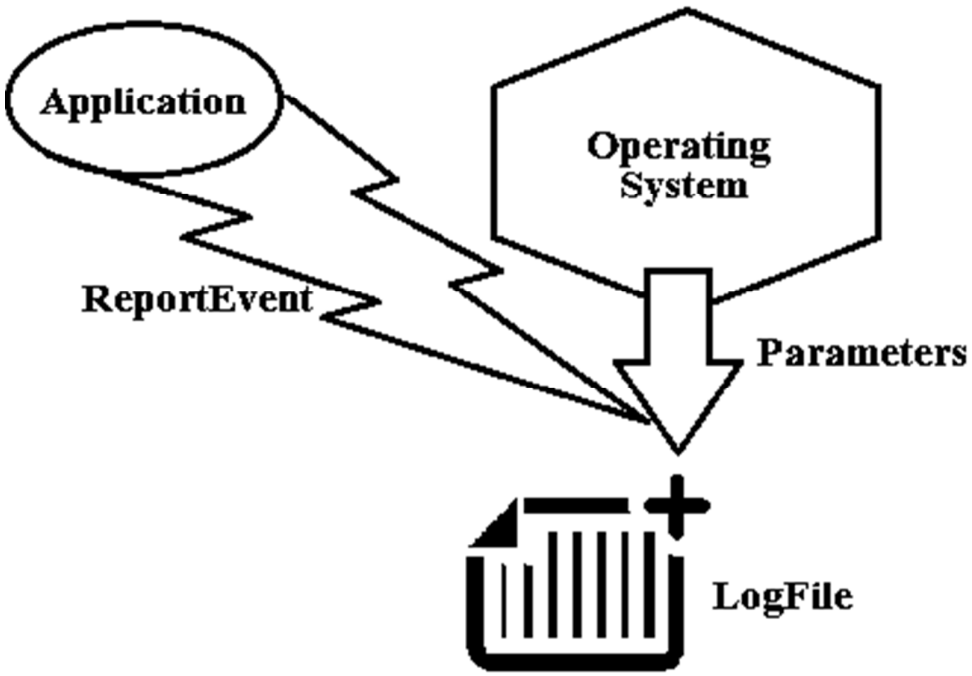

Figure 23: Event logs and reporting in windows

The event records are organized in either non-wrapping or wrapping way. The non-wrapping is a simple one where records are added between header and EOF record structures.

Non-wrapping:

HEADER (ELF_LOGFILE_HEADER)

EVENT 1	(EVENTLOGRECORD)
•	
•	
•	
EVENT 2	(EVENTLOGRECORD)
EOF RECORD	(ELF_EOF_RECORD)

Wrapping:

HEADER	(ELF_LOGFILE_HEADER)
PART OF EVENT N	(EVENTLOGRECORD)
EVENT N+1	(EVENTLOGRECORD)
•	
•	
•	
EOF RECORD	(ELF_EOF_RECORD)
Wasted space	
EVENT 1	(EVENTLOGRECORD)
EVENT 2	(EVENTLOGRECORD)
•	
•	
•	
PART OF EVENT N	(EVENTLOGRECORD)

The Wrapping mode uses circular way of adding new records. In this an old record is overwritten as new records come in.

2.4.2 Reading from an Windows event log file

On Windows the event logs can be managed with "Event Viewer" (eventvwr.msc) or "Windows Events Command Line Utility" (wevtutil.exe). Event Viewer can represent the EVTX (XML format) files in both "general view" (or formatted view) and "details view" (which has both a "friendly view" and "XML view"). Note that the formatted view can hide significant event data that is stored in the event record and can be seen in the detailed view.

An event viewer application like Windows Event Viewer or log parser uses the OpenEventLog function to open the event log for an event source. Then the viewer application uses the ReadEventLog function to read event records from the log. The following diagram illustrates this process (see figure 4).

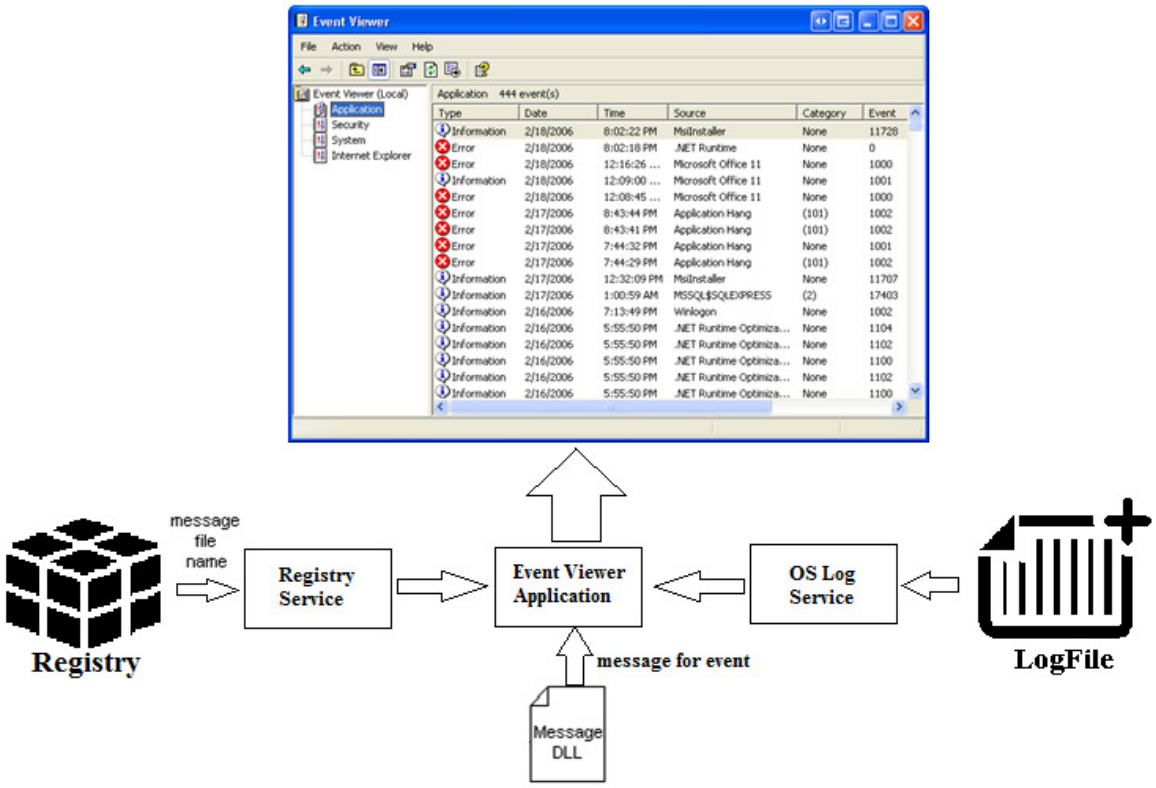

Figure 24: Process of viewing Event logs in windows.

2.4.3 Using Microsoft log parser

Logparser is a flexible command line utility that was initially written by Gabriele Giuseppini, a Microsoft employee, to automate tests for IIS logging. It was intended for use with the Windows operating system, and was included with the IIS 6.0 Resource Kit Tools. The default behavior of logparser works like a "data processing pipeline", by taking an SQL expression on the command line, and outputting the lines containing matches for the SQL expression.

Microsoft describes Logparser as a powerful, versatile tool that provides universal query access to text-based data such as log files, XML files and CSV files, as well as key data sources on the Windows operating system such as the Event Log, the Registry, the file system, and Active Directory. The results of the input query can be custom-formatted in text based output, or they can be persisted to more specialty targets like SQL, SYSLOG, or a chart.

Common usage:

$ logparser <options> <SQL expression>

Example: Selecting date, time and client username accessing ASPX-files, taken from all .log-files in the current directory.

$ logparser -i:IISW3C -q "SELECT date, time, cs-username FROM *.log WHERE cs-uri-stem LIKE '%.aspx' ORDER BY date, time;"

Figure 25: LogParser.

2.4.4 Understanding Windows user account management logs

Audit policies in windows can be edited using local group policy editor (see figure 2). Windows user account management security policy setting informs the operating system to logs when the following user account management tasks are performed:

- On creation, changing, deletion, renaming, disabling, enabling, locking out, or unlocking of an user account.
- On user account password change.
- On adding of Security identifier (SID) history to an user account.
- The restore mode password for Directory Services is set.
- Permissions on accounts are modified. Etc.

This policy setting is very useful for investigators in tracking events that involve getting sense of user accounts.

To view complete list of events in user account management please visit Microsoft site: *https://technet.microsoft.com/en-us/library/dn319091.aspx*.

2.4.5 Understanding Windows file and other object Access sets

Objects on internet or computer can be tracked using object access policy setting in audit events. If appropriate object access auditing subcategories (like file operations, Registry etc.) is enabled one can audit attempts to access a file, directory, registry key, or any other object (see figure 2). Many other subcategories are Audit Application Generated, Audit Certification Services, Audit Detailed File Share, Audit File Share, Audit File System, Audit Filtering Platform Connection, Audit Kernel Object, Audit Other Object Access Events, Audit Registry, Audit Security Account Management etc.

2.4.6 Auditing policy change

We can track the audit policy changes even. a local system or network Policy Change can be tracked using audit policy change events. Policies are mostly centrally created by admin or privileged users, thus, any changes or attempts to change these policies can be an important aspect of security management as well as while gathering investigative information. Few subcategories in this are: Audit Policy Change, Audit Authentication Policy Change, Audit Authorization Policy Change, Audit Filtering Platform Policy Change, Audit MPSSVC Rule-Level Policy Change, Audit Other Policy Change Events.

2.5 WINDOWS PASSWORD STORAGE

User and passwords in a window system are stored in either of two places:

a) SAM(Security Account Manager)
b) AD(Activity directory)

2.5.1 SAM

The Security Account Manager (SAM) is a database file in Windows XP, Windows Vista and Windows 7 that stores users' passwords. It can be used to authenticate local and remote users. Beginning with Windows 2000 SP4, Active Directory is used to authenticate remote users. SAM uses cryptographic measures to prevent forbidden users to gain access to the system.

The user passwords are stored in a hashed format in a registry hive either as a LM hash or as a NTLM hash. This file can be found in %SystemRoot%/system32/config/SAMand is mounted on HKLM/SAM.

In an attempt to improve the security of the SAM database against offline software cracking, Microsoft introduced the SYSKEY function in Windows NT 4.0. When SYSKEY is enabled, the on-disk copy of the SAM file is partially encrypted, so that the password hash values for all local accounts stored in the SAM are encrypted with a key (usually also referred to as the "SYSKEY"). It can be enabled by running the syskey program. Since a hash function is one-way, this provides some measure of security for the storage of the passwords.

In the case of online attacks, it is not possible to simply copy the SAM file to another location. The SAM file cannot be moved or copied while Windows is running, since the Windows kernel obtains and keeps an exclusive filesystem lock on the SAM file, and will not release that lock until the operating system has shut down or a "Blue Screen of Death" exception has been thrown. However, the in-memory copy of the contents of the SAM can be dumped using various techniques (including pwdump), making the password hashes available for offline brute-force attack.

2.5.1.1 Removing LM hash

Most versions of Windows can be configured to disable the creation and storage of valid LM hashes when the user changes their password. This is the default setting in Windows Vista, but was disabled by default in previous versions of Windows. Note: enabling this setting does not immediately clear the LM hash values from the SAM, but rather enables an additional check during password change operations that will instead store a "dummy" value in the location in the SAM database where the LM hash is otherwise stored. (This dummy value has no relationship to the user's password - it is the same value used for all user accounts.)

2.5.1.2 Related attacks

In Windows NT 3.51, NT 4.0 and 2000, an attack was devised to bypass the local authentication system. If the SAM file is deleted from the hard drive (e.g. mounting the Windows OS volume into an alternate operating system), the attacker could log in as any account with no password. This flaw was corrected with Windows XP, which shows an error message and shuts down the computer. However, there exist software utilities which, by the aforementioned methodology of using either an emulated virtual drive, or boot disk (usually Unix/Linux) based environment to mount the local drive housing the active NTFS partition, and using programmed software routines and function calls from within assigned memory stacks to isolate the SAM file from the Windows NT system installation directory structure (default: %SystemRoot%/system32/config/SAM) and, depending on the particular software utility being used, removes the password hashes stored for user accounts in their entirety, or in some cases, modify the user account passwords directly from this environment.

This software has both a highly pragmatic and beneficial use as a password clearing or account recovering utility for individuals who have lost or forgotten their windows account passwords, as well as a possible use as a malicious software security bypassing utility. Essentially granting a user with enough ability, experience, and familiarity with both the cracking utility software and the security routines of the Windows NT kernel (as well as offline and immediate local access to the target computer) the capability to entirely bypass/remove the windows account passwords from a potential target computer. Only recently, Microsoft released a utility called LockSmith, which is part of MSDart. MSDart is not freely available to end-users, however.

2.5.2 AD

Active Directory (AD) is a directory service that Microsoft developed for Windows domain networks and is included in most Windows Server operating systems as a set of processes and services.

An AD domain controller authenticates and authorizes all users and computers in a Windows domain type network—assigning and enforcing security policies for all computers and installing or updating software. For example, when a user logs into a computer that is part of a Windows domain, Active Directory checks the submitted password and determines whether the user is a system administrator or normal user.

Active Directory makes use of Lightweight Directory Access Protocol (LDAP) versions 2 and 3, Microsoft's version of Kerberos, and DNS.

As a directory service, an Active Directory instance consists of a database and corresponding executable code responsible for servicing requests and maintaining the database. The executable part, known as Directory System Agent, is a collection of Windows services and processes that run on Windows 2000 and later. Objects in Active Directory databases can be accessed via LDAP, ADSI (a component object model interface), messaging API and Security Accounts Manager services.

2.6 APPLICATION PASSWORDS CRACKERS

A password cracker is a program that can assist users to obtain unauthorised access to an application or resources. Also, Password crackers can help users to retrieve lost or forgotten passwords of any application.

2.6.1 Password cracking methods

Password crackers can use many ways to identify a password. The most important methods are:

a) Brute force method
b) Dictionary searches
c) Syllable attack
d) Rule based attack
e) Hybrid attack
f) Password guessing
g) Rainbow attack

2.6.1.1 Brute force attack

Brute force attacks work by calculating every possible combination that could make up a password and testing it to see if it is the correct password. As the password's length increases, the amount of time, on average, to find the correct password increases exponentially. This means short passwords can usually be discovered quite quickly, but longer passwords may take decades.

2.6.1.2 Dictionary attack

In cryptanalysis and computer security, a dictionary attack is a technique for defeating a cipher or authentication mechanism by trying to determine its decryption key or passphrase by trying hundreds or sometimes millions of likely possibilities, such as words in a dictionary.

A dictionary attack is based on trying all the strings in a pre-arranged listing, typically derived from a list of words such as in a dictionary (hence the phrase dictionary attack. In contrast to

a brute force attack, where a large proportion of the key space is searched systematically, a dictionary attack tries only those possibilities which are deemed most likely to succeed. Dictionary attacks often succeed because many people have a tendency to choose short passwords that are ordinary words or common passwords, or simple variants obtained, for example, by appending a digit or punctuation character. Dictionary attacks are relatively easy to defeat, e.g. by choosing a password that is not a simple variant of a word found in any dictionary or listing of commonly used passwords.

2.6.1.3 Syllable attack

It is a combination of the above two password attack. Many times the passwords does not contain a dictionary word and in these cases syllables form dictionary words use token and combined to every possible ways to do brute force searches.

2.6.1.4 Rule Based Attack

The attackers has many/ some preoccupied information using which the set of rules can be formed and then the possible searches can be narrowed down to a great extent. This type of attack is the most powerful one.

2.6.1.5 Hybrid attack and password guessing

It is also based on dictionary attack. In this if the old password is known than concatenating it with other symbols can yield the right password. In case of guessing the common passwords that are mostly used by novice users are used to crack codes.

2.6.1.6 Rainbow Attacks

Any computer system that requires password authentication must contain a database of passwords, either hashed or in plaintext, and various methods of password storage exist. Because the tables are vulnerable to theft, storing the plaintext password is dangerous. Most databases therefore store a cryptographic hash of a user's password in the database. In such a system, no one—including the authentication system—can determine what a user's password is simply by looking at the value stored in the database. Instead, when a user enters his or her password for authentication, it is hashed and that output is compared to the stored entry for that user (which was hashed before being stored). If the two hashes match, access is granted.

Someone who gains access to the (hashed) password table cannot merely enter the user's (hashed) database entry to gain access (using the hash as a password would of course fail since the authentication system would hash that a second time, producing a result which does not match the stored value, which was hashed only once). In order to learn a user's password, a password which produces the same hashed value must be found.

Rainbow tables are one tool that has been developed in an effort to derive a password by looking only at a hashed value.

Rainbow tables are not always needed, for there are simpler methods of hash reversal available. Brute-force attacks and dictionary attacks are the simplest methods available; however these are not adequate for systems that use large passwords, because of the difficulty

of storing all the options available and searching through such a large database to perform a reverse-lookup of a hash.

To address this issue of scale, reverse lookup tables were generated that stored only a smaller selection of hashes that when reversed could generate long chains of passwords. Although the reverse lookup of a hash in a chained table takes more computational time, the lookup table itself can be much smaller, so hashes of longer passwords can be stored. Rainbow tables are a refinement of this chaining technique and provide a solution to a problem called chain collisions.

A rainbow table is a pre-computed table for reversing cryptographic hash functions, usually for cracking password hashes. Tables are usually used in recovering a plaintext password up to a certain length consisting of a limited set of characters. It is a practical example of a space/time trade-off, using less computer processing time and more storage than a brute-force attack which calculates a hash on every attempt, but more processing time and less storage than a simple lookup table with one entry per hash. Use of a key derivation function that employs a salt makes this attack infeasible.

2.6.1.7 System passwords

Every personal computer and in that matter all computers does have a system setting or controls which are given by the hard core manufacturers to control access to system configuration and files that are vital to the boot process. In many cases users set password to the system control and there can be ways to break these passwords.

One way is to bypass the Bios password. Most of the manufacturers provide backup passwords. These can be accessed by reading their user documentation carefully. Like for example Dell gives backup password as “Dell” similarly Compaq gives as “Compaq”. However, if these backup passwords are not working one can use a combination of case sensitive back up passwords. Backup passwords are called as “Backdoor” passwords. While typing system passwords it should be known that typing wrong passwords can lock the entire system network and render a unstable device.

Another way is by re writing the CMOS Batter. Many times if the CMOS Battery is removed and replaced after 20-30 minutes the BIOS passwords resets.

Also, by adjusting the jumper settings on a mother board, all custom settings, including BIOS passwords will be cleaned. Location of these jumper settings may vary so we need to refer to the system documentation.

2.6.2 Tools for passwords cracking

There are several tools /software available to assist passwords recovery or cracking. Few examples are windows key generator, CMOSPwd, ERD commander.

2.6.2.1 CMOSPwd

CmosPwd decrypts password stored in cmos used to access BIOS SETUP. Works with the following BIOSes - ACER/IBM BIOS - AMI BIOS - AMI WinBIOS 2.5 - Award

4.5x/4.6x/6.0 - Compaq (1992) - Compaq (New version) - IBM (PS/2, Activa, Thinkpad) - Packard Bell - Phoenix 1.00.09.AC0 (1994), a486 1.03, 1.04, 1.10 A03, 4.05 rev 1.02.943, 4.06 rev 1.13.1107 - Phoenix 4 release 6 (User) - Gateway Solo - Phoenix 4.0 release 6 - Toshiba - Zenith AMI

2.6.2.2 ERDCommander

Microsoft DaRT is a successor of ERD Commander, which was part of the *Winternals Administrator Pack* from Winternals. ERD Commander later became a Microsoft property with its acquisition of Winternals on 17 July 2006.

Microsoft DaRT is based on Windows Preinstallation Environment now referred to as the Windows Recovery Environment. The tool set includes:

- Registry editor: Edits Windows Registry
- Locksmith: Resets a user account's password
- Crash Analyzer: Analyzes crash dumps
- File Restore: Restores deleted files
- Disk Commander: Repairs volumes, master boot records and partitions
- Disk Wipe: Irrecoverably erases data from hard disk
- Computer Management: A group of utilities that help retrieve system information, enable, disable or manage device drivers, Windows services and software that run during computer startup, inspect the event logs of the offline system and manage partitions.
- Explorer: A file manager
- Solution Wizard: A guidance tool that helps user choose the proper repair tool
- TCP/IP Config: Displays and modifies TCP/IP configuration
- Hotfix Uninstall: Uninstalls Windows hotfixes
- SFC Scan: Revives corrupted or deleted system files by copying them from the Windows installation source
- Search: Searches a disk for files
- Defender (formerly Standalone System Sweeper): An antivirus that scans a system for malware, rootkits, and potentially unwanted software. Uses the same engine as Microsoft Security Essentials and other Microsoft antivirus products.

ERD Commander originally included more tools, including a web browser.

2.6.2.3 Office pwd recovery

Office Password Recovery Toolbox is software which recovers lost password to any Microsoft Office document effectively. It can also recover read only files password. It allows several features to users letting them to set parameters to the searching password range like shape and length of the password. It enables users to search for string documents more efficiently and quickly. It recovers read only passwords from Microsoft Office Access. It is such type of application that can recover lost or forgotten password for Microsoft PowerPoint presentations, Microsoft Excel spreadsheets, Microsoft Access databases, Microsoft Outlook e-mail accounts, Microsoft OneNote notebooks etc. It can recover passwords instantly and helps in modifying sheet protection passwords, workbook passwords, email account

password, database passwords etc. It has user friendly interface which helps in extracting searches. The Office Password Recovery Tool provides an efficient access to MS Office documents.

Features:

- It recovers and removes all passwords of MS Excel, MS Outlook, MS Access documents, MS Word and VBA projects.
- It is able to crack all the Office document passwords and enables them for modifying workbook and worksheet passwords (Excel only), document protection passwords, database, user work group passwords and VBA project passwords.
- The entire recently opened password protected Microsoft Office documents is unprotected by using this software and opens the start up directly.
- It can access server's unique passwords and can break MS Excel or MS Word passwords irrespective of strength and length of password.
- It has the ability for protecting Office password Recovery Toolbox with password in order to prevent unauthorized access.

2.6.2.4 Passware kit

Passware Kit Enterprise and Forensics Passware Kit can recover the password of up to 150 different file types. It is trade, not exactly cheap tools, but can be very useful in different circumstances. This complete electronic evidence discovery solution reports all password-protected items on a computer and gains access to these items using the fastest decryption and password recovery algorithms. Many types of passwords are recovered or reset instantly, and advanced acceleration methods are used to recover difficult passwords. Passware Kit Forensic introduces a new attacks editor, which sets up the password recovery process in the most precise way to provide the quickest decryption solution possible. The highest performance is achieved with Distributed Password Recovery, using the computing power of multiple computers.

Passware Kit Forensic includes a Portable version that runs from a USB drive and finds encrypted files, recovers files and websites passwords without modifying files or settings on the host computer. Perform a complete encrypted evidence discovery process without installing Passware Kit on a target PC.

Passware Kit Forensic, complete with Passware FireWire Memory Imager, is the first commercial software that decrypts BitLocker and TrueCrypt hard disks of the seized computers without applying a time-consuming brute-force attack.

Key Features:

- Recovers passwords for 180+ file types and decrypts hard disks New!providing an all-in-one user interface
- Scans computers and network for password-protected files (Encryption Analyzer Professional included)
- Acquires memory images of the seized computers (FireWire Memory Imager included) New!

- Retrieves electronic evidence in a matter of minutes from a Windows Desktop Search Database (Search Index Examiner included)
- Supports Distributed and Cloud Computing password recovery New!
- Runs from a USB thumb drive and recovers passwords without installation on a target PC (Portable Version included)
- Includes 1-year Subscription to updates

General Features

- Instantly recovers many password types
- Instantly decrypts MS Word and Excel files up to version 2003 (20 Credits for Decryptum attack included)
- Resets passwords for Local and Domain Windows Administrators instantly
- Recovers encryption keys for hard disks protected with BitLocker in minutes, including BitLocker ToGo New!
- Decrypts TrueCrypt volumes in minutes New!
- Provides 8 different password recovery attacks (and any combination of them) with an easy-to-use setup wizard and drag & drop attacks editor
- Uses multiple-core CPUs and NVIDIA GPUs efficiently to speed up the password recovery process by up to 45 times
- Uses Tableau TACC hardware accelerators to speed up the password recovery process by up to 25 times
- Provides detailed reports with MD5 hash values

2.6.2.5 PDF Password Crackers

CrackPDF, Abcom PDF Password Cracker, and Advanced PDF Password Recovery can all be used to access password-protected Adobe PDF files. CrackPDF and Abcom PDF Password Cracker use brute force attacks to discover the passwords, while Advanced PDF Password Recovery simply removes the password protection entirely.

2.7 SUMMARY

1. Event logs and Password cracking plays very important role in digital forensics.
2. Event logging provides system administrators with information useful for diagnostics and auditing. Windows registry is also a very important source to maintain and manage logs.
3. Password cracking is utilized to gain access to digital evidence for which a judge has allowed access but the particular file's access is restricted.
4. Registry entries can be used to acquire and analyze much important information like system, time zone, shares, audit policy, wireless SSIDS, auto start locations, user login, activities, USB removable devices, trusted devices, cache, cookie and history etc.
5. User and passwords in a window system are stored in either Security Account Manager or Activity directory.

6. The most important methods of password cracking are brute force method, dictionary searches, syllable attack, rule based attack, hybrid attack, password guessing, rainbow attack.
7. There are several tools /software available to assist passwords recovery or cracking. Few examples are windows key generator, CMOSPwd, ERD commander.

2.8 CHECK YOUR PROGRESS

1. Fill in the blanks.

a) _____ and _____ can be very handy to get a good deal of analysis of registry entries.
b) When an application calls the _____ function to write an entry to the event log, the system passes the parameters to the _____.
c) An event viewer application uses the _____ function to open the event log for an event source.
d) SAM Stands for _____.
e) the _____ subcategory needs to be enabled to audit file operations and the _____ subcategory needs to be enabled to audit registry accesses

2. State True or False

a) In windows event logs are stored in binary format.
b) Policy Change audit events do not allow you to track changes to important security policies on a local system or network.
c) The Security Account Manager (SAM) is a database file in Windows.
d) Office Password Recovery Toolbox is software which stores lost password to any Microsoft Office document effectively.
e) Non-wrapping can occur when the event log is created or when the event log is cleared.

2.9 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) ProDiscover, ProScript
b) ReportEvent, event-logging service
c) OpenEventLog
d) Security Account Manager.
e) File System, Registry

2. State True or False

a) True
b) False
c) True
d) False
e) True.

2.10 FURTHER READINGS

1. Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
2. Investigating Hard Disks, File and Operating Systems: EC-Council | Press
3. Windows Event Log (EVT) – ForensicsWiki, www.forensicswiki.org/wiki/Windows_Event_Log_(EVT)
4. Audit User Account Management - TechNet – Microsoft, https://technet.microsoft.com/en-us/library/dd772693(v=ws.10).aspx
5. Event Log File Format (Windows) - MSDN – Microsoft, https://msdn.microsoft.com/en-us/library/windows/desktop/bb309026(v=vs.85).aspx
6. Policy Change - TechNet – Microsoft, https://technet.microsoft.com/en-us/library/dd772669(v=ws.10).aspx
7. Reading from the Event Log (Windows) - MSDN – Microsoft, https://msdn.microsoft.com/en-us/library/windows/desktop/aa363675(v=vs.85).aspx

References, Article Source & Contributors

[1] Active Directory - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Active_Directory

[2] CMOSPwd, https://packages.gentoo.org/packages/app-forensics/cmospwd

[3] Dictionary attack - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Dictionary_attack

[4] Event logging – Wikipedia, https://en.wikipedia.org/wiki/Event_logging

[5] Log analysis - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Log_analysis \

[6] logparser - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Logparser

[7] Microsoft Desktop Optimization Pack - Wikipedia, , https://en.m.wikipedia.org/.../Microsoft_Diagnostics_and_Recovery_Tool

[8] Passware kit, http://azizalstsetia.blogspot.in/2011/04/passware-kit-forensic-103-full-version_7549.html

[9] Password cracking - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Password_cracking.

[10] Rainbow table - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Rainbow_table

[11] Recover lost MS Office Password, recoverlostofficepassword.wikidot.com

[12] Security Account Manager - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Security_Account_Manager

[13] Windows XML Event Log, (EVTX), http://www.forensicswiki.org/wiki/Windows_XML_Event_Log_(EVTX)

UNIT III: NETWORK FORENSICS

3.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand basic concepts of networking and its role in forensics.
- Understand the OSI and TCP/IP Layers and basic protocols which are pertinent for forensics.
- Understand intrusion detection and prevention systems.
- Understand ways of capturing of network logs.
- Understand ways analysing network time stamps and data logs.
- Know and understand usage of various network tools used in forensics.
- Know and understand usage of various software tools used in forensics.

3.2 INTRODUCTION

There are many systems that track and record network activities and data. However, there are still some measures that add up to the forensics on network systems. The network forensics adds vital information to investigations. Tools can be used to do time line analysis, email re-construction, Metadata analysis, packet frame analysis or checksum on data exchanged.

Another aspect of network forensics is to make/ get capabilities of capturing and investigating a suspect's computer over network. There are methods of making an image of a suspect/ victims computer over network connection form the forensics lab itself. However, legal aspects must be considered before capturing/ intruding over other system. Network forensics is a sub-branch of digital forensics relating to the monitoring and analysis of computer network traffic for the purposes of information gathering, legal evidence, or intrusion detection. Unlike other areas of digital forensics, network investigations deal with volatile and dynamic information. Network traffic is transmitted and then lost, so network forensics is often a pro-active investigation. Network forensics generally has two uses. The first, relating to security, involves monitoring a network for anomalous traffic and identifying intrusions. An attacker might be able to erase all log files on a compromised host; network-based evidence might therefore be the only evidence available for forensic analysis.[3] The second form relates to law enforcement. In this case analysis of captured network traffic can include tasks such as reassembling transferred files, searching for keywords and parsing human communication such as emails or chat sessions. In 2000 the FBI lured computer hackers Aleksey Ivanov and Gorshkov to the United States for a fake job interview. By monitoring network traffic from the pair's computers, the FBI identified passwords allowing them to collect evidence directly from Russian-based computers.

3.3 NETWORK COMPONENTS AND THEIR FORENSICS IMPORTANCE

A computer network or data network is a telecommunications network which allows computers to exchange data. In computer networks, networked computing devices exchange data with each other along network links (data connections). The connections between nodes

are established using either cable media or wireless media. The best-known computer network is the Internet.

Network computer devices that originate, route and terminate the data are called network nodes. Nodes can include hosts such as personal computers, phones, servers as well as networking hardware. Two such devices can be said to be networked together when one device is able to exchange information with the other device, whether or not they have a direct connection to each other.

Computer networks differ in the transmission media used to carry their signals, the communications protocols to organize network traffic, the network's size, topology and organizational intent. In most cases, communications protocols are layered on (i.e. work using) other more specific or more general communications protocols, except for the physical layer that directly deals with the transmission media. Computer networks support applications such as access to the World Wide Web, shared use of application and storage servers, printers, and fax machines, and use of email and instant messaging applications.

3.3.1 Host

A network host is a computer or other device connected to a computer network. A network host may offer information resources, services, and applications to users or other nodes on the network. A network host is a network node that is assigned a network layer host address.

Computers participating in networks that use the Internet Protocol Suite may also be called IP hosts. Specifically, computers participating in the Internet are called Internet hosts, sometimes Internet nodes. Internet hosts and other IP hosts have one or more IP addresses assigned to their network interfaces. The addresses are configured either manually by an administrator, automatically at start-up by means of the Dynamic Host Configuration Protocol (DHCP), or by stateless address auto-configuration methods.

Every network host is a physical network node (i.e. a network device), but not every physical network node is a host. Network devices such as modems, hubs and network switches are not assigned host addresses (except sometimes for administrative purposes), and are consequently not considered to be network hosts. Devices such as network printers and hardware routers have IP addresses, but since they are not general-purpose computers, they are sometimes not considered to be hosts.

Network hosts that participate in applications that use the client-server model of computing, are classified as server or client systems. Network hosts may also function as nodes in peer-to-peer applications, in which all nodes share and consume resources in an equipotent manner.

In operating systems, the term terminal host traditionally denotes a multi-user computer or software providing services to computer terminals, or a computer that provides services to smaller or less capable devices, such as a mainframe computer serving teletype terminals or video terminals. Other examples are a telnet host (a telnet server) and an xhost (X Window client).

3.3.2 Node

In data communication, a physical network node may either be a data communication equipment (DCE) such as a modem, hub, bridge or switch; or a data terminal equipment(DTE) such as a digital telephone handset, a printer or a host computer, for example a router, a workstation or a server.

If the network in question is a LAN or WAN, every LAN or WAN nodes (that are at least data link layer devices) must have a MAC address, typically one for each network interface controller it possesses. Examples are computers, packet switches, xDSL modems (with Ethernet interface) and wireless LAN access points. Note that a hub constitutes a physical network node, but does not constitute a LAN network node, since a hubbed network logically is a bus network. Analogously, a repeater or PSTN modem (with serial interface) is a physical network node but not a LAN node in this sense.

If the network in question is the Internet or an Intranet, many physical network nodes are host computers, also known as Internet nodes, identified by an IP address, and all hosts are physical network nodes. However, some datalink layer devices such as switches, bridges and WLAN access points do not have an IP host address (except sometimes for administrative purposes), and are not considered to be Internet nodes or hosts, but as physical network nodes and LAN nodes.

If the network in question is a distributed system, the nodes are clients, servers or peers. A peer may sometimes serve as client, sometimes server. In a peer-to-peer or overlay network, nodes that actively route data for the other networked devices as well as themselves are called super nodes.

Distributed systems may sometimes use virtual nodes so that the system is not oblivious to the heterogeneity of the nodes. This issue is addressed with special algorithms, like consistent hashing, as it is the case in Amazon's.

3.3.3 Router

A router is a networking device that forwards data packets between computer networks. Routers perform the "traffic directing" functions on the Internet. A data packet is typically forwarded from one router to another through the networks that constitute the internetwork until it reaches its destination node.

A router is connected to two or more data lines from different networks (as opposed to a network switch, which connects data lines from one single network). When a data packet comes in on one of the lines, the router reads the address information in the packet to determine its ultimate destination. Then, using information in its routing table or routing policy, it directs the packet to the next network on its journey. This creates an overlay internetwork.

The most familiar type of routers are home and small office routers that simply pass data, such as web pages, email, IM, and videos between the home computers and the Internet. An

example of a router would be the owner's cable or DSL router, which connects to the Internet through an ISP. More sophisticated routers, such as enterprise routers, connect large business or ISP networks up to the powerful core routers that forward data at high speed along the optical fiber lines of the Internet backbone. Though routers are typically dedicated hardware devices, use of software-based routers has grown increasingly common.

3.3.4 Switch

A network switch (also called switching hub, bridging hub, officially MAC Bridge) is a computer networking device that connects devices together on a computer network, by using packet switching to receive, process and forward data to the destination device. Unlike less advanced network hubs, a network switch forwards data only to one or multiple devices that need to receive it, rather than broadcasting the same data out of each of its ports.

A network switch is a multiport network bridge that uses hardware addresses to process and forward data at the data link layer (layer 2) of the OSI model. Switches can also process data at the network layer (layer 3) by additionally incorporating routing functionality that most commonly uses IP addresses to perform packet forwarding; such switches are commonly known as layer-3 switches or multilayer switches.

A switch is a device in a computer network that electrically and logically connects together other devices. Multiple data cables are plugged into a switch to enable communication between different networked devices. Switches manage the flow of data across a network by transmitting a received message only to the one or more devices for which the message was intended. Each networked device connected to a switch can be identified using a MAC address, allowing the switch to regulate the flow of traffic. This maximizes the security and efficiency of the network.

3.3.5 Hub

An Ethernet hub, active hub, network hub, repeater hub, multiport repeater, or simply hub is a device for connecting multiple Ethernet devices together and making them act as a single network segment. It has multiple input/output (I/O) ports, in which a signal introduced at the input of any port appears at the output of every port except the original incoming. A hub works at the physical layer (layer 1) of the OSI model. Repeater hubs also participate in collision detection, forwarding a jam signal to all ports if it detects a collision. In addition to standard 8P8C ("RJ45") ports, some hubs may also come with a BNC or Attachment Unit Interface (AUI) connector to allow connection to legacy 10BASE2 or 10BASE5 network segments.

Hubs are now largely obsolete, having been replaced by network switches except in very old installations or specialized applications.

3.3.6 Network interface card (NIC)

A network interface controller (NIC, also known as a network interface card, network adapter, LAN adapter or physical network interface, and by similar terms) is a computer hardware component that connects a computer to a computer network. A device that usually

holds the MAC (Media Access Control) address of your computer that uniquely identifies your host or computer. The NIC is the physical bridge between the network and the host. If you see on the back of your computer a wire with an oversized phone jack and blinking lights, it is NIC.

3.4 OSI

The Open Systems Interconnection model (OSI Model) is a conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard to their underlying internal structure and technology. Its goal is the interoperability of diverse communication systems with standard protocols. The model partitions a communication system into abstraction layers. The original version of the model defined seven layers.

A layer serves the layer above it and is served by the layer below it. For example, a layer that provides error-free communications across a network provides the path needed by applications above it, while it calls the next lower layer to send and receive packets that comprise the contents of that path. Two instances at the same layer are visualized as connected by a horizontal connection in that layer.

3.4.1 OSI model

The OSI (Open System Interconnection) is a standard logical view of any networking. It has 7 layers as given in *Figure 26*. *Figure 26* also depicts various form of data formats that are exchanged between each layers of either side in a connectivity. It also gives main functionalities of each layer as abstracted.

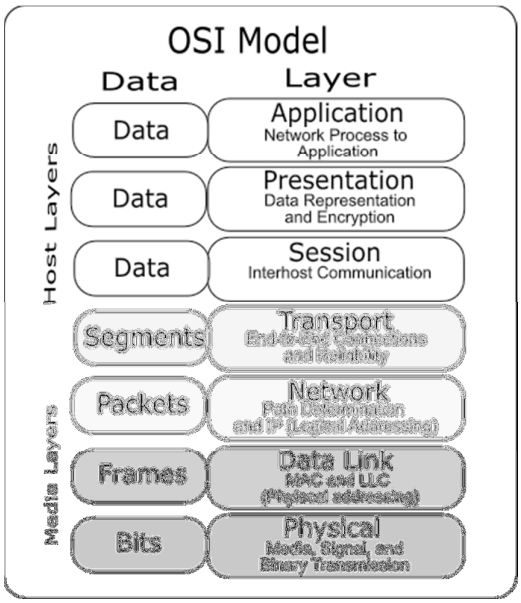

Figure 26: OSI Layers

3.4.2 TCP/IP Layers

Although the OSI model is widely used and often cited as the standard, TCP/IP protocol has been used by most UNIX workstation vendors. TCP/IP is designed around a simple four-layer

scheme. It does omit some features found under the OSI model. Also it combines the features of some adjacent OSI layers and splits other layers apart. The four network layers defined by TCP/IP model are as follows (also given in the figure 2):

- Layer 1 – Link: This layer defines the network hardware and device drivers.
- Layer 2 – Network: This layer is used for basic communication, addressing and routing. TCP/IP uses IP and ICMP protocols at the network layer.
- Layer 3 – Transport: Handles communication among programs on a network. TCP and UDP fall within this layer.
- Layer 4 – Application: End-user applications reside at this layer. Commonly used applications include NFS, DNS, arp, rlogin, talk, ftp, ntp and traceroute.

The Internet protocol suite is the computer networking model and set of communications protocols used on the Internet and similar computer networks. It is commonly known as TCP/IP, from Transmission Control Protocol (TCP) and the Internet Protocol (IP).

TCP/IP provides end-to-end connectivity specifying how data should be packetized, addressed, transmitted, routed and received at the destination. This functionality is organized into four abstraction layers which are used to sort all related protocols according to the scope of networking involved. From lowest to highest, the layers are the link layer, containing communication technologies for a single network segment (link); the internet layer, connecting hosts across independent networks, thus establishing internetworking; the transport layer handling host-to-host communication; and the application layer, which provides process-to-process application data exchange.

The TCP/IP model and related protocol models are maintained by the Internet Engineering Task Force (IETF).

Encapsulation is used to provide abstraction of protocols and services. Encapsulation is usually aligned with the division of the protocol suite into layers of general functionality. In general, an application (the highest level of the model) uses a set of protocols to send its data down the layers, being further encapsulated at each level.

The layers of the protocol suite near the top are logically closer to the user application, while those near the bottom are logically closer to the physical transmission of the data. Viewing

layers as providing or consuming a service is a method of abstraction to isolate upper layer protocols from the details of transmitting bits over, for example, Ethernet and collision detection, while the lower layers avoid having to know the details of each and every application and its protocol.

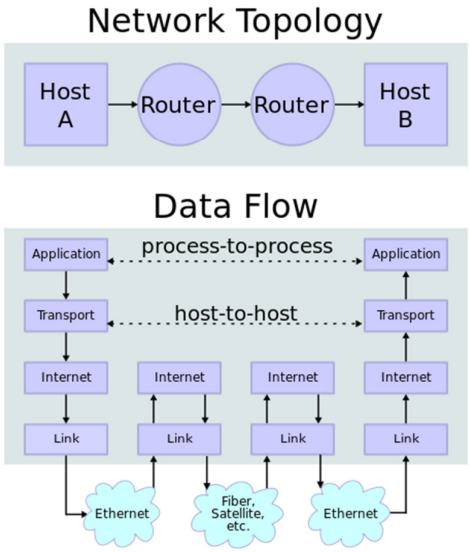

Figure 27:Internetworking.

Figure 27 depicts two Internet hosts connected via two routers and the corresponding layers used at each hop. The application on each host executes read and write operations as if the processes were directly connected to each other by some kind of data pipe. Every other detail of the communication is hidden from each process. The underlying mechanisms that transmit data between the host computers are located in the lower protocol layers.

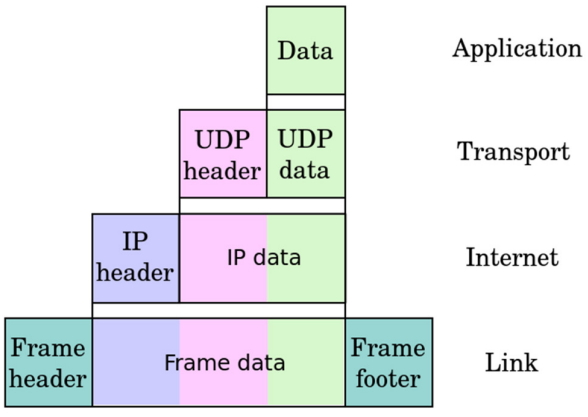

Figure 28: Encapsulation of application data descending through the layers

Applying forensic methods on the Ethernet layer is done by eavesdropping bit streams with tools called monitoring tools or sniffers. The most common tool on this layer is Wireshark (formerly known as Ethereal) and tcpdump where tcpdump works mostly on unix-like operating systems. These tools collect all data on this layer and allow the user to filter for different events. With these tools, websites, email attachments, and other network traffic can be reconstructed only if they are transmitted or received unencrypted. An advantage of collecting this data is that it is directly connected to a host. If, for example the IP address or the MAC address of a host at a certain time is known, all data sent to or from this IP or MAC address can be filtered.

To establish the connection between IP and MAC address, it is useful to take a closer look at auxiliary network protocols. The Address Resolution Protocol (ARP) tables list the MAC addresses with the corresponding IP addresses.

To collect data on this layer, the network interface card (NIC) of a host can be put into "promiscuous mode". In so doing, all traffic will be passed to the CPU, not only the traffic meant for the host.

However, if an intruder or attacker is aware that his connection might be eavesdropped, he might use encryption to secure his connection. It is almost impossible to break nowadays encryption but the fact that a suspect's connection to another host is all the time encrypted might indicate that the other host is an accomplice of the suspect.

On the network layer the Internet Protocol (IP) is responsible for directing the packets generated by TCP through the network (e.g., the Internet) by adding source and destination information which can be interpreted by routers all over the network. Cellular digital packet networks, like GPRS, use similar protocols like IP, so the methods described for IP work with them as well.

For the correct routing, every intermediate router must have a routing table to know where to send the packet next. These routing tables are one of the best sources of information if investigating a digital crime and trying to track down an attacker. To do this, it is necessary to follow the packets of the attacker, reverse the sending route and find the computer the packet came from (i.e., the attacker).

The internet can be a rich source of digital evidence including web browsing, email, newsgroup, synchronous chat and peer-to-peer traffic. For example web server logs can be used to show when (or if) suspect accessed information related to criminal activity. Email accounts can often contain useful evidence; but email headers are easily faked and, so, network forensics may be used to prove the exact origin of incriminating material. Network forensics can also be used in order to find out who is using a particular computer by extracting user account information from the network traffic.

3.5 FORENSICS INFORMATION FROM NETWORK

Major information sources in network are: Host, router, fireworks, switches, and intrusion detection and prevention systems, network printers/copiers etc. wireless access points. An investigator needs to collect data from these sources. The categorization of these data as well as way these needs to be collected and analysed is of utmost importance.

Hosts: Generally forensics makes use of agents (Software) to gather and send Host data to remote forensic server. The agents collect real time data stream passing through the network interface card (NIC) and send for analysis study.

Routers: Mostly router logs can be useful in many cases. Information of status details, errors, IP and MAC addresses getting resolved to other networks or hosts can be used to trace a suspect as well as can be helpful in getting to the chain of events while restructuring the crime.

Firewalls: Firewalls also very importantly maintain logs of every internet/ network access by the host user. These logs can be like dropped packets, un allowed application, filtered websites, recognised attacks, etc. at many times the logs of the host firewall or the network firewall is enough to trace the logs of the host firewall ir the network firewall is enough to trace links to a crime or suspicious activity.

Switch: Switches have a CAM (context addressable memory) which keeps information about mappings of MAC address to ports. Also, CAM is used to keep information about VLAN.

Two popular methods that are specifically designed to allow a network analyst to monitor traffic are [10]:

1. Port mirroring – the switch sends a copy of network packets to a monitoring network connection.
2. SMON – "Switch Monitoring" is described by RFC 2613 and is a protocol for controlling facilities such as port mirroring.

3.5.1 Intrusion detection/ prevention system

An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a management station. IDS come in a variety of "flavors" and approach the goal of detecting suspicious traffic in different ways. There are network based (NIDS) and host based (HIDS) intrusion detection systems. NIDS is a network security system focusing on the attacks that come from the inside of the network (authorized users). Some systems may attempt to stop an intrusion attempt but this is neither required nor expected of a monitoring system. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, and reporting attempts.

Though they both relate to network security, an intrusion detection system (IDS) differs from a firewall in that a firewall looks outwardly for intrusions in order to stop them from happening. Firewalls limit access between networks to prevent intrusion and do not signal an

attack from inside the network. An IDS evaluates a suspected intrusion once it has taken place and signals an alarm. An IDS also watches for attacks that originate from within a system.

The logs generated by the IDS can be very useful for network forensics analysis.

Certain times network printers/copiers etc. also log the activities to some extent and can play vital role in network forensics. However, the logs maintained depend upon the manufacturer.

3.5.2 Wireless Access Points

At times WAP can also come into play as it also maintains logs of almost all routing type activities that it does like SSIDs and incoming connections etc. It is to be noted, looking at the amount of traffic that follows in and out of a network it is important to understand the storage aspects also. That is, how we will be storing these logs etc. for future analysis as well as evidence building.

The investigators can use one or more of the available bilk storage technologies like SAN (storage area network), network attached storage (NAS), direct attached storage (DAS) etc. for the purpose. Also, tape drives are in use since older days and still play a vital role in mass storages.

3.6 LOG ANALYSIS

The analysis of large volumes of data collected during IDPS is typically performed in a separate database system run by the analysis team. Live systems are usually not dimensioned to run extensive individual analysis without affecting the regular users. On the other hand, it is methodically preferable to analyze data copies on separate systems and protect the analysis teams against the accusation of altering original data.

Due to the nature of the data, the analysis focuses more often on the content of data than on the database it is contained in. If the database itself is of interest then Database forensics are applied.

In order to analyze large structured data sets with the intention of detecting financial crime it takes at least three types of expertise in the team: A data analyst to perform the technical steps and write the queries, a team member with extensive experience of the processes and internal controls in the relevant area of the investigated company and a forensic scientist who is familiar with patterns of fraudulent behavior.

After an initial analysis phase using methods of explorative data analysis the following phase is usually highly iterative. Starting with a hypothesis on how the perpetrator might have created a personal advantage the data is analyzed for supporting evidence. Following that the hypothesis is refined or discarded.

The combination of different databases, in particular data from different systems or sources is highly effective. These data sources are either unknown to the perpetrator or he/she cannot manipulate them afterwards. Data Visualization is often used to display the results.

There are many tools that can be used to analyse the logs captured during above sources of information. However, still we need to understand how these analysis are done and how actually a criminal event can be re-created. Major activities during log analysis are:

a) Analysing time stamps
b) Analysing data

3.6.1 Analyzing time stamps

Time and its synchronization in network are very important. A smart criminal can use certain methodologies to put false time stamps in their communication. However with advent of technologies like Network Time Protocol (NTP) this issue is more or less minimized. The investigator needs to find out whether the NTP has been incorporated or not before proceeding into the analysis. Network Time Protocol (NTP) is a networking protocol for clock synchronization between computer systems over packet-switched, variable-latency data networks. In operation since before 1985, NTP is one of the oldest Internet protocols in current use. NTP was originally designed by David L. Mills of the University of Delaware, who still oversees its development. NTP is intended to synchronize all participating computers to within a few milliseconds of Coordinated Universal Time (UTC).

3.6.2 Analyzing data

Data over network in Transmission Control Protocol/ Internet Protocol (TCP/IP) is broken into pieces which are further broken into smaller pieces called as packets to be transported over networks. The packets are re-assembled at the other end. Different packets of the same message might take different paths before reaching at other end. This adds to the complexity of reassembling the packets. To overcome this issue TCP/IP follows a mechanism of numbering each packet based on sequences. The receiver node sends acknowledgment based on these sequence numbers. The message is reconstructed and the sending host gets acknowledgement of all the packets sent over the network. The times stamps in these acknowledgement packets are in GMT (UTC) format and can give vital clues during analysis.

Other protocol which has to be understood are Address resolution protocol (ARP) which is used to map MAC address to an IP and vis-versa. This resolution protocols can help an investigator get vital traces into IP addresses and MAC addresses of any individual in a case. Other protocols/ technologies that need an overview are Internet control message protocol (ICMP), Internet protocol security (IPSec), BitTorrent, Domain name system (DNS), Dynamic host configuration protocol (DHCP), File transfer protocol (FTP), HyperText Transfer Protocol (HTTP), Internet message access protocol (IMAP), Network time protocol (NTP), Post office protocol version 3 (POP3), Secure shell (SSH), Simple mail transfer protocol (SMTP) etc.

3.7 FORENSICS TOOLS

Forensic tools that are used for forensic activities like seizure, capture, analysis etc. in network can be categorized in two forms:

a. Technology tools
b. Software tools

Technology tools are like methodologies to track, trace or identify hidden artefacts in any network system. The software tools are software solutions which can specifically assist forensic collection etc.

3.7.1 Network tools used for forensics

Network tap

A network tap is a hardware device which provides a way to access the data flowing across a computer network. In many cases, it is desirable for a third party to monitor the traffic between two points in the network. If the network between points A and B consists of a physical cable, a "network tap" may be the best way to accomplish this monitoring. The network tap has (at least) three ports: an A port, a B port, and a monitor port. A tap inserted between A and B passes all traffic through unimpeded, but also copies that same data to its monitor port, enabling a third party to listen.

Network taps are commonly used for network intrusion detection systems, VoIP recording, network probes, RMON probes, packet sniffers, and other monitoring and collection devices and software that require access to a network segment. Taps are used in security applications because they are non-obtrusive, are not detectable on the network (having no physical or logical address), can deal with full-duplex and non-shared networks, and will usually *pass through* traffic even if the tap stops working or loses power.

Once a network tap is in place, the network can be monitored without interfering with the network itself. Other network monitoring solutions require in-band changes to network devices, which meant that monitoring can impact the devices being monitored. Once a tap is in place, a monitoring device can be connected to it as-needed without impacting the monitored network.

Putting a network tap into place will disrupt the network being monitored for a short time. Even so, a short disruption is preferable to taking a network down multiple times to deploy a monitoring tool. Establishing good guidelines for the placement of network taps is recommended.

Port Mirroring

Port mirroring is used on a network switch to send a copy of network packets seen on one switch port (or an entire VLAN) to a network monitoring connection on another switch port.

Promiscous mode

In computer networking, promiscuous mode (often shortened to "promisc mode" or "promisc. mode") is a mode for a wired network interface controller (NIC) or wireless network interface controller (WNIC) that causes the controller to pass all traffic it receives to the central processing unit (CPU) rather than passing only the frames that the controller is intended to receive. This mode is normally used for packet sniffing that takes place on a router or on a computer connected to a hub (instead of a switch) or one being part of a WLAN. Interfaces are placed into promiscuous mode by software bridges often used with hardware virtualization.

Promiscuous mode is often used to diagnose network connectivity issues. There are programs that make use of this feature to show the user all the data being transferred over the network. Some protocols like FTP and Telnet transfer data and passwords in clear text, without encryption, and network scanners can see this data. Therefore, computer users are encouraged to stay away from insecure protocols like telnet and use more secure ones such as SSH.

3.7.2 Software tools used for network forensics

Wire shark

Wireshark is a free and open-source packet analyzer. It is used for network troubleshooting, analysis, software and communications protocol development, and education. Originally named Ethereal, the project was renamed Wireshark in May 2006 due to trademark issues.

Wireshark is cross-platform, using the GTK+ widget toolkit in current releases, and Qt in the development version, to implement its user interface, and using pcap to capture packets; it runs on Linux, OS X, BSD, Solaris, some other Unix-like operating systems, and Microsoft Windows. There is also a terminal-based (non-GUI) version called TShark. Wireshark, and the other programs distributed with it such as TShark, are free software, released under the terms of the GNU General Public. Wireshark is very similar to tcpdump, but has a graphical front-end, plus some integrated sorting and filtering options. Figure 4 depicts a typical wireshark gui.

Wireshark lets the user put network interface controllers that support promiscuous mode into that mode, so they can see all traffic visible on that interface, not just traffic addressed to one of the interface's configured addresses and broadcast/multicast traffic. However, when capturing with a packet analyzer in promiscuous mode on a port on a network switch, not all traffic through the switch is necessarily sent to the port where the capture is done, so capturing in promiscuous mode is not necessarily sufficient to see all network traffic. Port mirroring or various network taps extend capture to any point on the network.

Wireshark is software that "understands" the structure (encapsulation) of different networking protocols. It can parse and display the fields, along with their meanings as specified by different networking protocols. Wireshark uses pcap to capture packets, so it can only capture packets on the types of networks that pcap supports.

- Data can be captured "from the wire" from a live network connection or read from a file of already-captured packets.
- Live data can be read from a number of types of networks, including Ethernet, IEEE 802.11, PPP, and loopback.
- Captured network data can be browsed via a GUI, or via the terminal (command line) version of the utility, TShark.
- Captured files can be programmatically edited or converted via command-line switches to the "editcap" program.
- Data display can be refined using a display filter.
- Plug-ins can be created for dissecting new protocols.

- VoIP calls in the captured traffic can be detected. If encoded in a compatible encoding, the media flow can even be played.
- Raw USB traffic can be captured.

Wireshark's native network trace file format is the libpcap format supported by libpcap and WinPcap, so it can exchange captured network traces with other applications that use the same format, including tcpdump and CA NetMaster. It can also read captures from other network analyzers, such as snoop, Network General's Sniffer, and Microsoft Network Monitor.

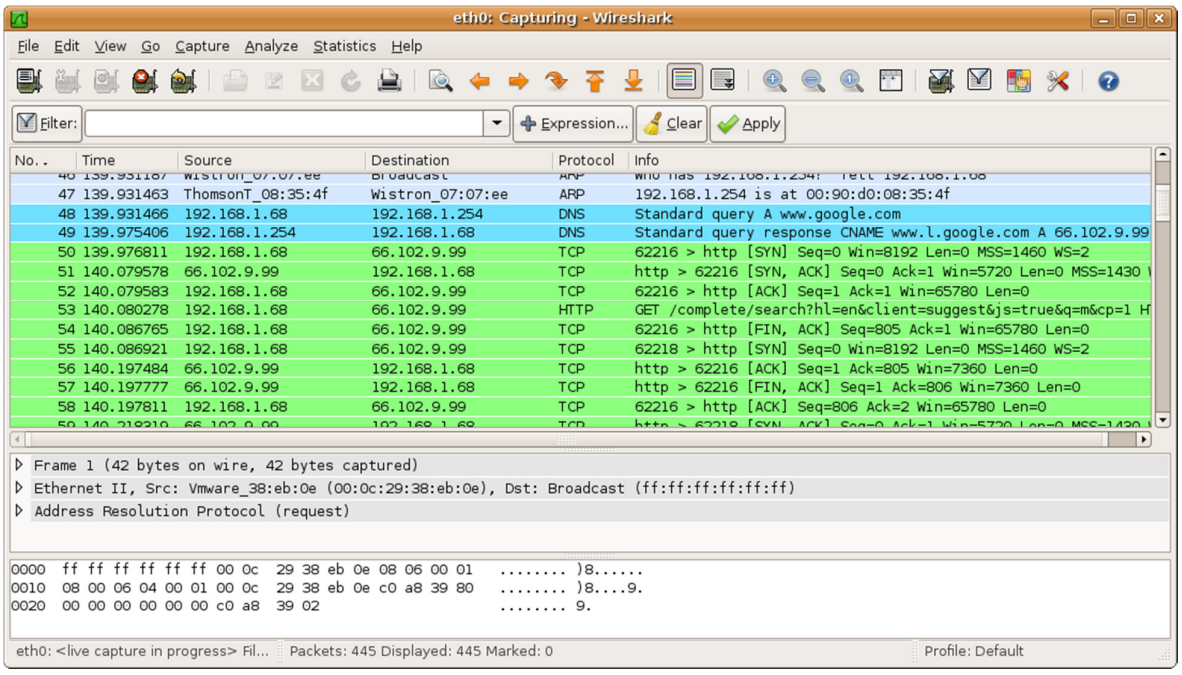
eth0: Capturing - Wireshark
File Edit View Go Capture Analyze Statistics Help
Filter:
Expression... Clear Apply
No. Time Source Destination Protocol Info
47 139.931463 ThomsonT_08:35:4f Wistron_07:07:ee ARP 192.168.1.254 is at 00:90:d0:08:35:4f
48 139.931466 192.168.1.68 192.168.1.254 DNS Standard query A www.google.com
49 139.975406 192.168.1.254 192.168.1.68 DNS Standard query response CNAME www.l.google.com A 66.102.9.99
50 139.976811 192.168.1.68 66.102.9.99 TCP 62216 > http [SYN] Seq=0 Win=8192 Len=0 MSS=1460 WS=2
51 140.079578 66.102.9.99 192.168.1.68 TCP http > 62216 [SYN, ACK] Seq=0 Ack=1 Win=5720 Len=0 MSS=1430
52 140.079583 192.168.1.68 66.102.9.99 TCP 62216 > http [ACK] Seq=1 Ack=1 Win=65780 Len=0
53 140.080278 192.168.1.68 66.102.9.99 HTTP GET /complete/search?hl=en&client=suggest&js=true&q=m&cp=1
54 140.086765 192.168.1.68 66.102.9.99 TCP 62216 > http [FIN, ACK] Seq=805 Ack=1 Win=65780 Len=0
55 140.086921 192.168.1.68 66.102.9.99 TCP 62218 > http [SYN] Seq=0 Win=8192 Len=0 MSS=1460 WS=2
56 140.197484 66.102.9.99 192.168.1.68 TCP http > 62216 [ACK] Seq=1 Ack=805 Win=7360 Len=0
57 140.197777 66.102.9.99 192.168.1.68 TCP http > 62216 [FIN, ACK] Seq=1 Ack=806 Win=7360 Len=0
58 140.197811 192.168.1.68 66.102.9.99 TCP 62216 > http [ACK] Seq=806 Ack=2 Win=65780 Len=0
Frame 1 (42 bytes on wire, 42 bytes captured)
Ethernet II, Src: Vmware_38:eb:0e (00:0c:29:38:eb:0e), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
Address Resolution Protocol (request)
0000 ff ff ff ff ff ff 00 0c 29 38 eb 0e 08 06 00 01 ....)8......
0010 08 00 06 04 00 01 00 0c 29 38 eb 0e c0 a8 39 80 ....)8....9.
0020 00 00 00 00 00 00 c0 a8 39 02 .... 9.
eth0: <live capture in progress> Fil... Packets: 445 Displayed: 445 Marked: 0 Profile: Default

Figure 29: Wireshark GUI

TCPDUMP

Tcpdump is a common packet analyser that runs under the command line. It allows the user to display TCP/IP and other packets being transmitted or received over a network to which the computer is attached. Distributed under the BSD license, tcpdump is free software.

Tcpdump works on most Unix-like operating systems: Linux, Solaris, BSD, OS X, HP-UX, Android and AIX among others. In those systems, tcpdump uses the libpcap library to capture packets. The port of tcpdump for Windows is called WinDump; it uses WinPcap, the Windows port of libpcap.

Tcpdump prints the contents of network packets. It can read packets from a network interface card or from a previously created saved packet file. Tcpdump can write packets to standard output or a file.

It is also possible to use tcpdump for the specific purpose of intercepting and displaying the communications of another user or computer. A user with the necessary privileges on a system acting as a router or gateway through which unencrypted traffic such

as Telnet or HTTP passes can use tcpdump to view login IDs, passwords, the URLs and content of websites being viewed, or any other unencrypted information.

3.8 SUMMARY

1. Network forensics is a sub-branch of digital forensics relating to the monitoring and analysis of computer network traffic for the purposes of information gathering, legal evidence, or intrusion detection.

2. Network components like host, node, router, switch, hub, NIC etc. all have to be considered while examining a network forensically.

3. OSI and TCP/IP layers needs to be understood while doing forensics over networks.

4. Applying forensic methods on the Ethernet layer is done by eavesdropping bit streams with tools called monitoring tools or sniffers like wireshark and tcpdump.

5. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, and reporting attempts. The logs generated by the IDS can be very useful for network forensics analysis.
6. In order to analyse large structured data sets with the intention of detecting financial crime it takes at least three types of expertise in the team: A data analyst to perform the technical steps and write the queries, a team member with extensive experience of the processes and internal controls in the relevant area of the investigated company and a forensic scientist who is familiar with patterns of fraudulent behaviour.
7. There are many tools that can be used to analyse the time stamps as well as data of the logs captured during Intrusion detection and prevention systems and monitoring above sources of information (components) in a network.
8. Technology tools are like methodologies to track, trace or identify hidden artefacts in any network system. The software tools are software solutions which can specifically assist forensic collection etc.
9. Tools can be used to do time line analysis, email re-construction, Metadata analysis, packet frame analysis or checksum on data exchanged.

3.9 CHECK YOUR PROGRESS

1. Fill in the blanks.

i. Computers participating in networks that use the Internet Protocol Suite may also be called _____.
ii. Modem, hub, bridge or switches are _____ in a data communication.
iii. Digital telephone handset, a printer or a host computer are called as _____ in a data communication.

iv. A _____ is a networking device that forwards data packets between computer networks. Routers perform the _____ functions on the Internet.

v. A _____ is a computer networking device that connects devices together on a computer network, by using packet switching to receive, process and forward data to the destination device.

vi. A _____ is a computer hardware component that connects a computer to a computer network.

vii. TCP/IP model has basically 4 layers they are: _____, _____, _____, _____ Layers.

viii. Applying forensic methods on the Ethernet layer is done by eavesdropping bit streams with tools called _____.

ix. A switch sends a copy of network packets to a monitoring network connection is called as _____.

x. _____ are primarily focused on identifying possible incidents, logging information about them, and reporting attempts.

2. State True or False

i. Generally forensics makes use of agents (Software) to gather and send Host data to remote forensic server.

ii. Routers have a CAM (context addressable memory) which keeps information about mappings of MAC address to ports.

iii. Firewalls are example of IDS.

iv. Network Time Protocol (NTP) is a networking protocol for clock synchronization between computer systems over packet-switched, variable-latency data networks.

3.10 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) IP hosts.
b) data communication equipment (DCE).
c) data terminal equipment(DTE.
d) router , "traffic directing".
e) network switch .
f) A network interface controller .
g) Link, Network,Transport, Application.
h) monitoring tools or sniffers.
i) Port Mirroring.
j) Intrusion detection and prevention systems (IDPS).

2. State True or False

i. (T)
ii. (F)
iii. (F)
iv. (T)

3.11 FURTHER READINGS

1. Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
2. Investigating Hard Disks, File and Operating Systems: EC-Council | Press
3. Gary Palmer, A Road Map for Digital Forensic Research, Report from DFRWS 2001, First Digital Forensic Research Workshop, Utica, New York, August 7 – 8, 2001, Page(s) 27–30

3.12 MODEL QUESTIONS

1. State and explain various network components and their forensic importance.
2. How are the network logs captured and analysed? Explain.
3. What are IDS and IDPS?
4. State major features of wireshark tool.
5. What is promiscuous mode in networking?
6. What do you understand be network tapping and port mirroring?

References, Article Source & Contributors

[1] Computer network - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Computer_network

[2] Ethernet hub - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Ethernet_hub

[3] Forensic data analysis - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Forensic_data_analysis

[4] Host (network) - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Host_(network

[5] Intrusion detection system - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Intrusion_detection_system

[6] Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.

[7] Network forensics - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Network_forensics

[8] Network interface controller - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Network_interface_controller

[9] Network switch - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Network_switch

[10] Network tap - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Network_tap

[11] Network Time Protocol - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Network_Time_Protocol

[12] Node (networking) - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Node_(networking)

[13] OSI model - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/OSI_model

[14] Port mirroring - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Port_mirroring
[15] Promiscuous mode - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Promiscuous_mode
[16] Router (computing) - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Router_(computing)
[17] TCP/IP 4 layer model, http://www.planetlarg.net/tcpip-4-layer-model
[18] tcpdump - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Tcpdump
[19] Wireshark - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Wireshark

UNIT IV: WIRELESS ATTACKS

4.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand wireless networking.
- Understand frames and their capturing for forensics.
- Understand various attacks in wireless networking and correlate to the forensics.
- Understand wireless intrusion detection techniques using available tools.

4.2 INTRODUCTION

Wireless forensics is a sub-discipline of network forensics. The main goal of wireless forensics is to provide the methodology and tools required to collect and analyze (wireless) network traffic that can be presented as valid digital evidence in a court of law. The evidence collected can correspond to plain data or, with the broad usage of Voice-over-IP (VoIP) technologies, especially over wireless, can include voice conversations. Analysis of wireless network traffic is similar to that on wired networks; however there may be the added consideration of wireless security measures. Wireless networks have entered in a paramount way in the day to day life of people as well as enterprises. The wireless have added convenience of mobility and thus introduced risks on the traditional networks.

We will first look into wireless technologies (mainly 802.11) through the wireless frame layer (OSI Layer) and understand how they can be captured, extracted and analysed. After that we will learn various wireless attacks and the intrusion detection systems in wireless layers.

4.3WIRELESS FIDELTY (WI-FI)(802.11)

The Wi-Fi Alliance defines Wi-Fi as any "wireless local area network" (WLAN) product based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards.[1] However, the term "Wi-Fi" is used in general English as a synonym for "WLAN" since most modern WLANs are based on these standards. "Wi-Fi" is a trademark of the Wi-Fi Alliance. Many devices can use Wi-Fi, e.g. personal computers, video-game consoles, smartphones, digital cameras, tablet computers and digital audio players. These can connect to a network resource such as the Internet via a wireless network access point. Such an access point (or hotspot) has a range of about 20 meters (66 feet) indoors and a greater range outdoors. Hotspot coverage can be as small as a single room with walls that block radio waves, or as large as many square kilometres achieved by using multiple overlapping access points.

Wi-Fi provides service in private homes, businesses, as well as in public spaces at Wi-Fi hotspots set up either free-of-charge or commercially, often using a captive portal webpage for access. Organizations and businesses, such as airports, hotels, and restaurants, often provide free-use hotspots to attract customers. Enthusiasts or authorities who wish to provide services or even to promote business in selected areas sometimes provide free Wi-Fi access. A service set is the set of all the devices associated with a particular Wi-Fi network. The service set can be local, independent, extended or mesh. Each service set has an associated identifier, the Service Set Identifier (SSID), which consists of 32 bytes that identifies the particular network. The SSID is configured within the devices that are considered part of the network,

and it is transmitted in the packets. Receivers ignore wireless packets from other networks with a different SSID.

The 802.11 logical architecture consists of several components (see *Figure 30*): station (STA), wireless access point (AP), independent basic service set (IBSS), basic service set (BSS), distribution system (DS), and extended service set (ESS). STAs and Aps are hardware devices. The wireless STA has an adapter card, PC Card, or an embedded device to facilitate wireless connectivity. The wireless AP provides access to wireless STAs by becoming a bridge between STAs and the existing network backbone for network access.

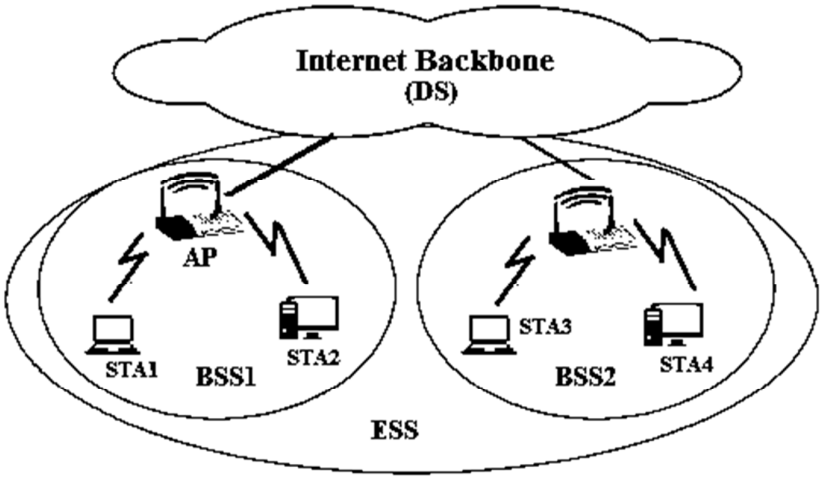

Figure 30: 802.11 components.

4.3.1 Capturing 802.11 frames

Air being the media, data packets is open for anyone to be sniffed. The sniffer setup must be equipped with appropriate hardware and software. Various ways of sniffing into wireless frames are using capabilities of monitor mode, using software like kismet and packet analysers etc.

Monitor mode

Monitor mode, or RFMON (Radio Frequency MONitor) mode, allows a computer with a wireless network interface controller (WNIC) to monitor all traffic received from the wireless network. Unlike promiscuous mode, which is also used for packet sniffing, monitor mode allows packets to be captured without having to associate with an access point ad hoc network first. Monitor mode only applies to wireless networks, while promiscuous mode can be used on both wired and wireless networks. Monitor mode is one of the seven modes that 802.11 wireless cards can operate in: Master (acting as an access point), Managed (client, also known as station), Ad hoc, Mesh, Repeater, Promiscuous, and Monitor mode.

Software such as KisMAC or Kismet, in combination with packet analysers that can read pcap files, provide a user interface for passive wireless network monitoring. In many cases, monitor mode support is not properly implemented by the vendor. Linux's interfaces

for 802.11 drivers support monitor mode and many drivers offer that support. FreeBSD, NetBSD, OpenBSD, and DragonFly BSD also provide an interface for 802.11 drivers that supports monitor mode and many drivers for those operating systems support monitor mode as well.

Kismet

Kismet is a network detector, packet sniffer, and intrusion detection system for 802.11 wireless LANs. Kismet will work with any wireless card which supports raw monitoring mode, and can sniff 802.11a, 802.11b, 802.11g, and 802.11n traffic. The program runs under Linux, FreeBSD, NetBSD, OpenBSD, and Mac OS X. The client can also run on Microsoft Windows, although, aside from external drones (see *Figure 31*), there's only one supported wireless hardware available as packet source. Distributed under the GNU General Public License, Kismet is free software.

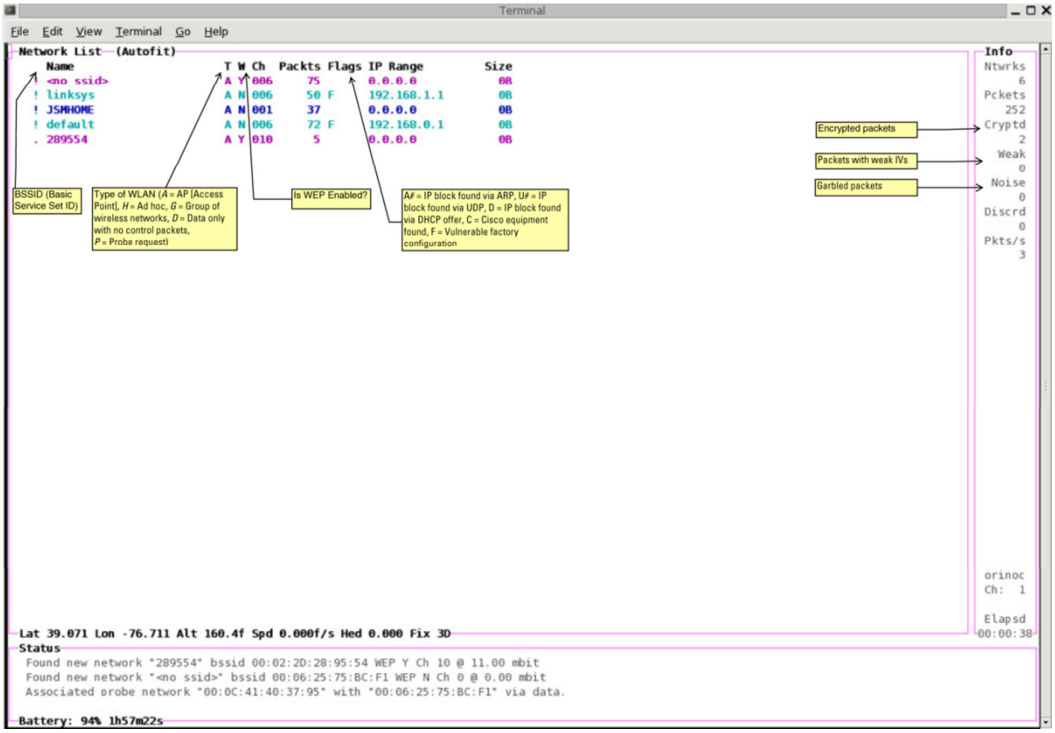
Terminal
File Edit View Terminal Go Help
Network List (Autofit)
Name T W Ch Packts Flags IP Range Size
<no ssid> A Y 006 75 0.0.0.0 0B
! linksys A N 006 50 F 192.168.1.1 0B
! JSMHOME A N 001 37 0.0.0.0 0B
! default A N 006 72 F 192.168.0.1 0B
. 289554 A Y 010 5 0.0.0.0 0B
Info
Ntwrks 6
Pckets 252
Cryptd 2
Weak 0
Noise 0
Discrd 0
Pkts/s 3
Encrypted packets
Packets with weak IVs
Garbled packets
BSSID (Basic Service Set ID)
Type of WLAN (A = AP [Access Point], H = Ad hoc, G = Group of wireless networks, D = Data only with no control packets, P = Probe request)
Is WEP Enabled?
A# = IP block found via ARP, U# = IP block found via UDP, D = IP block found via DHCP offer, C = Cisco equipment found, F = Vulnerable factory configuration
orinoc
Ch: 1
Elapsd
00:00:38
Lat 39.071 Lon -76.711 Alt 160.4f Spd 0.000f/s Hed 0.000 Fix 3D
Status
Found new network "289554" bssid 00:02:2D:28:95:54 WEP Y Ch 10 @ 11.00 mbit
Found new network "<no ssid>" bssid 00:06:25:75:BC:F1 WEP N Ch 0 @ 0.00 mbit
Associated probe network "00:0C:41:40:37:95" with "00:06:25:75:BC:F1" via data.
Battery: 94% 1h57m22s

Figure 31: An explanation of the headings displayed in Kismet.

Kismet in tandem with wireshark can be used to capture and analyse wireless packets. Major attributes that needs to be located and further analysed are: wireless packets, basic system ID, Frame Sequence number, WEP etc. Packets captured by Kismet can be saved into pcap files, which are then analysed by some analyser tools by opening those files in an offline mode.

NetStumbler

NetStumbler (also known as Network Stumbler) is a tool for Windows that facilitates detection of Wireless LANs using the 802.11b, 802.11a and 802.11g WLAN standards. It runs on Microsoft Windows operating systems from Windows 2000 to Windows XP. A

trimmed-down version called MiniStumbler is available for the handheld Windows CE operating system. No updated version has been developed since 2005.

The program is commonly used for:

i. Wardriving: Wardriving is the act of searching for Wi-Fi wireless networks by a person in a moving vehicle, using a portable computer, smartphone or personal digital assistant (PDA).
ii. Verifying network configurations
iii. Finding locations with poor coverage in a WLAN
iv. Detecting causes of wireless interference
v. Detecting unauthorized ("rogue") access points
vi. Aiming directional antennas for long-haul WLAN links. (A directional antenna or beam antenna is an antenna which radiates or receives greater power in specific directions allowing for increased performance and reduced interference from unwanted sources.)

Pcap

In the field of computer network administration, pcap (packet capture) consists of an application programming interface (API) for capturing network traffic. Unix-like systems implement pcap in the libpcap library; Windows uses a port of libpcap known asWinPcap.

Monitoring software may use libpcap and/or WinPcap to capture packets travelling over a network and, in newer versions, to transmit packets on a network at the link layer, as well as to get a list of network interfaces for possible use with libpcap or WinPcap.

Airodump and aircrack

Airodump-ng is a Packet sniffer, it Places air traffic into PCAP or IVS files and shows information about networks. Aircrack-ng is a network software suite consisting of a detector, packet sniffer, WEP and WPA/WPA2-PSK cracker and analysis tool for 802.11 wireless LANs. It works with any wireless network interface controller whose driver supports raw monitoring mode and can sniff 802.11a, 802.11b and 802.11g traffic. The program runs under Linux and Windows.

WEPWedgie

WEPWedgie is a open source toolkit for determining 802.11 WEP keystreams and injecting traffic with known keystreams. The toolkit also includes logic for firewall rule mapping, pingscanning, and portscanning via the injection channel and a cellular modem

4.4 WIRELESS SECURITY

Wireless security is the prevention of unauthorized access or damage to computers using wireless networks. The most common types of wireless security are Wired Equivalent Privacy (WEP) and Wi-Fi Protected Access (WPA). WEP is a notoriously weak security standard. The password it uses can often be cracked in a few minutes with a basic laptop computer and widely available software tools. WEP is an old IEEE 802.11 standard from 1999, which was outdated in 2003 by WPA, or Wi-Fi Protected Access. WPA was a quick

alternative to improve security over WEP. The current standard is WPA2; some hardware cannot support WPA2 without firmware upgrade or replacement. WPA2 uses an encryption device that encrypts the network with a 256-bit key; the longer key length improves security over WEP. Short for Wired Equivalent Privacy (or Wireless Encryption Protocol), WEP is part of the IEEE 802.11 wireless networking standard and was designed to provide the same level of security as that of a wired LAN. Because wireless networks broadcast messages using radio, they are susceptible to eavesdropping. WEP provides security by encrypting data over radio waves so that it is protected as it is transmitted from one end point to another. WEP was the encryption scheme considered to be the initial standard for first generation wireless networking devices. However, it has been found that WEP is not as secure as once believed.

4.4.1 Attacks

Wireless systems encounter attacks which are in some cases similar to network attacks. However, the attacks which are typically specific to wireless systems are:

a) Probing and surveillance.
b) Denial of Service.
c) Impersonation or MAC Spoofing.
d) Man in the middle.

Probing and surveillance

Probing or sniffing can be of two types:

a) Active
b) Passive

Attackers can indulge in active probing where they send probe requests and continuously wait for a probe response. The response will contain SSID information and many other information from nodes or access points in the range. Certain access points are cloaked, i.e. they are configured not to respond with a probe request. In such cases the attacker might not get any active response hence will not be able to probe or sniff into these access points.

In passive probing the attacker keeps on listening on all available (or listenable) channels for all the packets that are sent or received. While doing this the attacker doesn't have to send a single packet into the transmission channel. But, cloaked Access Points with no wireless activities during the period of the probe would not be detected. Because there is no probes the cloaked Access Points will not send any packets into the channel, thus, the attacker will never be able to sniff into those Access Points.

NetStumbler is a good example of a tool that can help in active probing is. Kismet is a software tool that facilitates passive probing. The Data gathered during probing can be saved into pcap format (see previous section) for future analysis while in offline mode. On a non-encrypted stream in the network, the attacker could immediately find or probe into a traffic stream and can easily acquire vital information MAC address, IP address range, and gateway etc from the traffic.

In case of encrypted streams like Wireless Encryption Protocol (WEP), then WEP crackers which are available with the attacker can be used. For example, airodump can be used to

gather all the encrypted packets transmitted and aircrack (see previous section) is then used to try to crack the WEP key. If there is no sufficient traffic on the network, certain tools for packet injection like WEPWedgie (see previous section) can be used to insert random traffic into the WEP encrypted network. This will fetch responses from the network; these response packets can be collected and given for WEP key cracking.

Denial of Service

DoS type attacks at every wireless layers can be easily carried out in a wireless network. Noise I the channel can be induced by emitting a very strong Radio Frequency interference on the channel in which the wireless network is operating on this will cause interference to all wireless networks that are operating on that channel or nearby channels. Certain DoS attacks can utilize packet injection, the attackers will flood the network connected clients with lots of disassociate or authentication packets.

Impersonation(spoofing)

Another attack is called as impersonation, where the attackers change their MAC address in the transmission packets with an address that he had found while probing. This is typically used by criminals to send derogatory mails like intimidation etc. A MAC address might belong to an authorized client in the network. This is usually done to defeat the MAC filtering capabilities of access points where only a list of authorized MAC addresses are allowed to use the wireless network. As earlier described, even if the wireless network is WEP encrypted, the MAC address of the sending and receiving party is still viewable by a wireless sniffing tool. MAC spoofing is a technique for changing a factory-assigned Media Access Control (MAC) address of a network interface on a networked device. The MAC address is hard-coded on a network interface controller (NIC) and cannot be changed. However, there are tools which can make an operating system believe that the NIC has the MAC address of a user's choosing. The process of masking a MAC address is known as MAC spoofing. Essentially, MAC spoofing entails changing a computer's identity, for any reason, and it is relatively easy. MAC address can be changed in linux using ifconfig command. In windows we can do this using registry.

Man in the middle

A man-in-the-middle attacker entices computers to log into a computer which is set up as a soft AP (Access Point). Once this is done, the hacker connects to a real access point through another wireless card offering a steady flow of traffic through the transparent hacking computer to the real network. The hacker can then sniff the traffic. One type of man-in-the-middle attack relies on security faults in challenge and handshake protocols to execute a “de-authentication attack”. This attack forces AP-connected computers to drop their connections and reconnect with the hacker’s soft AP (disconnects the user from the modem so they have to connect again using their password which one can extract from the recording of the event). Man-in-the-middle attacks are enhanced by software such as LANjack and AirJack which automate multiple steps of the process, meaning what once required some skill can now be done by script kiddies. Hotspots are particularly vulnerable to any attack since there is little to no security on these networks.

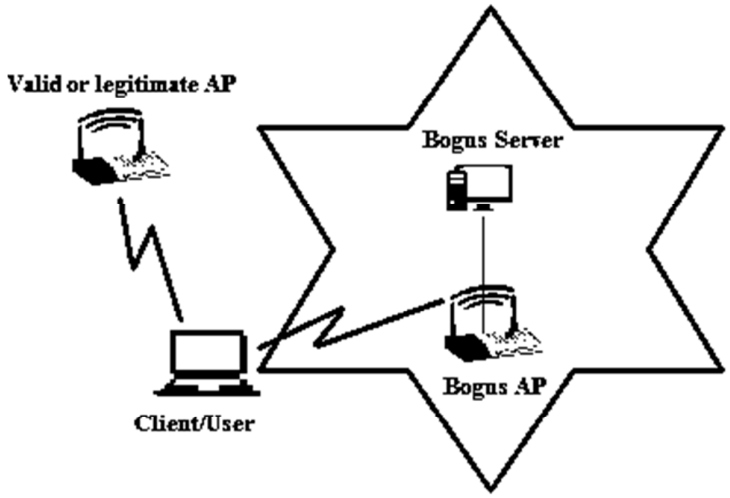

Figure 32: A typical man in middle in wireless systems

4.5 WIRELESS ATTACKS DETECTION TECHNIQUES

Now that we have a good idea of various attacks in a wireless system, we should now look into certain ways that can be employed to detect certain attacks. These detection techniques can be categorized in following three basic forms:

a. Wireless Access point monitoring
b. Wireless client/node monitoring
c. Wireless traffic monitoring

4.5.1 Wireless Access Point Monitoring

In this the wireless network keeps a list of authorized access points and hardware using the net with information like respective SSID, MAC address and other channel information recorded earlier. The monitoring agent/component would continuously listen to wireless frames like beacons, frame probes; responses and authentications etc. sent out by every Access Points and compare these with the previously recorded information. The monitoring device must listen to every possible channel and record all packets for this technique to be effective. To detect Man-in-the-middle attack, such a monitoring component needs to detect that whether there is a sudden introduction of an AP on another channel previously not present. Though the SSID, MAC address might be spoofed (see previous section) by the attacker in the process of setting up the rouge AP, the channel information in which the genuine AP was operating from has been changed which provides an alert on a possible MitM attack.

4.5.2 Wireless Client/Node Monitoring

The access point monitoring is much simpler, in the wireless client monitoring a list of allowed clients' needs to be maintained. This adds up to lot of administrative overheads, however, some of the clients aspects can be recorded and monitored. Like, list of blacklisted clients can be maintained and any movements from these nodes can generate alerts for analysis. Also, all wireless clients with an unauthorized MAC address (MAC address ranges

which have not been allocated out yet) are automatically denied access and an alert send off. Also, clients sending probes with typical nicknames can also be recorded and alert generated. One more area where monitoring might be applied is WEP (encrypted) traffic is being used to send/receive, no station should be reusing the same WEP Initialization Vector (used to generate keys) over and over again within a very short period of time (WepWedgie and other cracking tools use this).

For wireless clients that are legitimate, there is a sequence number field within the IEEE 802.11 header which can be tracked for abrupt changes. Certain times when impersonation attacks are being carried out, the attacker will be able to read the MAC / IP address of the victim, but it will not be able to continue with the sequence number used previously by the victim, thus by monitoring the sequence number in these client generated packets impersonation attacks can be easily detected.

4.5.3 General Wireless Traffic Monitoring

To detect DoS attacks, Wireless traffic can be monitored for attempts to flood the network using deauthentication, de-association, authentication, association, erroneous authentication. Frequency and Signal-To-Noise Ratio monitoring could help signal an oncoming RF based DOS attack on your wireless network. Failures in authentication as well as association can also be monitored and reported.

4.6 WIRELESS INTRUSION DETECTION SYSTEMS

Let us look at few examples of open source wireless Intrusion Detection Systems that are available for usage.

4.6.1 Snort-wireless

Snort's open source network-based intrusion detection system (NIDS) has the ability to perform real-time traffic analysis and packet logging on Internet Protocol (IP) networks.

Snort can be configured in three main modes: sniffer, packet logger, and network intrusion detection. In sniffer mode, the program will read network packets and display them on the console. In packet logger mode, the program will log packets to the disk. In intrusion detection mode, the program will monitor network traffic and analyse it against a rule set defined by the user. The program will then perform a specific action based on what has been identified. Snort-wireless is a wireless intrusion detection system adapted from the snort IDS engine. One can write snort-wireless rules for detecting wireless traffic like one would detect for IP layer attacks.

4.6.2 WIDZ

WIDZ version 1 is a proof of concept IDS system for 802.11 that guards an AP(s) and Monitors local frequencies for potentially malevolent activity. It detects scans, association floods, and bogus/Rogue AP's. It can easily be integrated with SNORT or RealSecure.

4.6.3 Bro

Originally written by Vern Paxson, Bro is an open source Unix based network monitoring framework. Often compared to a Network intrusion detection systems (NIDS), Bro can be used to build a NIDS but is much more. Bro can also be used for collecting network measurements, conducting forensic investigations, traffic baselining and more. Bro has been compared to tcpdump, Snort, netflow, and Perl (or any other scripting language) all in one. It is released under the BSD license.

Bro can be conceptualized in two layers

Bro Event Engine; which analyses live or recorded network traffic or trace files to generate neutral events.

Bro uses an engine (written in C++) to generate events when "something" happens. This can be triggered by the Bro process, such as just after initialization or just before termination of the Bro process, as well as by something taking place on the network (or trace file) being analysed, such as Bro witnessing an HTTP request or a new TCP connection. Bro uses common ports and dynamic protocol detection (involving signatures as well as behavioural analysis) to make a best guess at interpreting network protocols. Events are policy neutral in that they are not good or bad but simply signals to script land that something happened.

Bro Policy Scripts; which analyse events to create action policies.

Events are handled from within Bro policy scripts (written in the Turing complete Bro scripting language). By default Bro simply logs information about events to files (Bro also supports logging events in binary output), however it can be configured to take other actions such as sending an email, raising an alert, executing a system command, updating an internal metric and even calling another Bro script. The default behaviour produces net flow-like output (conn log) as well as application event information. Bro scripts are able to read in data from external files, such as blacklists, for use within Bro policy scripts.

4.7 SUMMARY

1. Wireless networks have entered in a paramount way in the day to day life of people as well as enterprises. The wireless have added convenience of mobility and thus introduced risks on the traditional networks.
2. The IEEE 802.11 protocol and associated technologies are the basis of present day wireless networking.
3. Various ways of sniffing into wireless frames are by using capabilities of monitor mode.
4. WEP is an old IEEE 802.11 standard from 1999, which was outdated in 2003 by WPA, or Wi-Fi Protected Access. WPA was a quick alternative to improve security over WEP. The current standard is WPA2.
5. Attacks which are typically specific to wireless systems are Probing and surveillance, Denial of Service, Impersonation or MAC Spoofing, Man in the middle.

6. Wireless attack detection techniques can be categorized in following three basic forms; Wireless Access point monitoring, Wireless client/node monitoring, Wireless traffic monitoring.
7. Few examples of open source wireless Intrusion Detection Systems that are available for usage are Snort-wireless, WIDZ, RealSecure.

4.8 CHECK YOUR PROGRESS

1. Fill in the blanks.

a) Main components in the 802.11 are _____.
b) Various ways of sniffing into wireless frames are using capabilities of _____.
c) _____ is the act of searching for Wi-Fi wireless networks by a person in a moving vehicle, using a portable computer, smartphone or personal digital assistant (PDA).
d) WEP stands for _____.
e) WPA stands for _____.
f) Examples of wireless intrusion detection systems are: _____.
g) Two ways of DoS attack in wireless systems are _____, _____.

2. State True or False

a) Monitor mode is one way of capturing packets and applies to both wired and wireless networks.
b) pcap (packet capture) consists of an application programming interface (API) for capturing network traffic.
c) WEP enabled node is highly secured.
d) WPA stands for Wireless Protection and authentication.
e) Active probing is where an attacker sends probe requests and continuously wait for a probe response from an access point.
f) Impersonation is to use captured MAC address while communicating.
g) Wireless Access Point Monitoring helps in detecting spoofing and man in the middle attacks.
h) In Wireless Client/Node Monitoring the administrator continuously sends probe packets to clients connected to an access point.

4.9 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) Station (STA), wireless access point (AP), independent basic service set (IBSS), basic service set (BSS), distribution system (DS), and extended service set (ESS).
b) Monitor mode.
c) Wardriving.
d) Wired Equivalent Privacy.
e) Wi-Fi Protected Access.
f) Snort-wireless, WIDZ, RealSecure.

g) Inducing strong RF noise, continuously injecting lot of authentication packets.

2. State True or False.

a) (F)
b) (T)
c) (F)
d) (F)
e) (T)
f) (T)
g) (T)
h) (F)

4.10 FURTHER READING

a) Debra Littlejohn Shinder, Michael Cross, Scene of the Cybercrime, syngress
b) Computer Forensics: Investigating Wireless Networks and Devices By EC-Council
c) Mark Ciampa, CWSP Guide to Wireless Security
d) Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
e) How 802.11 Wireless Works: Wireless - TechNet – Microsoft, https://technet.microsoft.com/en-us/library/cc757419(v=ws.10).aspx
f) Intrusion Detection Systems: An Overview of RealSecure, https://www.sans.org/reading-room/whitepapers/detection/intrusion-detection-systems-overview-realsecure-342
g) Understanding Wireless Attacks and Detection – SANS, https://www.sans.org/.../understanding-wireless-attacks-detection-1633

References, Article Source & Contributors

[1] Aircrack-ng - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Aircrack-ng
[2] Kismet (software) - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Kismet_(software)
[3] Monitor mode - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Monitor_mode
[4] NetStumbler - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/NetStumbler
[5] Network forensics - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Network_forensics
[6] pcap - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Pcap
[7] The Differences Between WEP and WPA - Webopedia.com, www.webopedia.com › Did You Know › Computer_Science
[8] WEPWedgie - Best Open Source, www.findbestopensource.com/product/wepwedgie
[9] widzv1-0.zip ≈ Packet Storm, https://packetstormsecurity.com/files/30700/widzv1-0.zip.html

[10] Wireless security - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Wireless_security
[11] Bro (software) - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Bro_(software)

BLOCK III

UNIT I: INVESTIGATING WEB ATTACKS

1.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand cyber-attacks and cyber warfare.
- Know to categorize specific cyber-attacks.
- Know basic forensics involved in cyber-attacks investigation.
- Know various applications or software used during web forensics.

1.2 INTRODUCTION

The web application plays a very important in today's people's lives. The first generation of web application were limited by static HTML applications. Later on the internet and web access became quite ubiquitous and the user's expectations from web application also increased many folds. Common gateway Interface (CGI) provided a big leap in the direction of modern web applications. The users were facilitated with more features like, searching, hosting uploading etc.The CGI provided more interactive forms over internet for the users to interact. Newer more advanced frameworks came into foray like, PHP, ASP.NET, J2EE, AJAX, Ruby on Rails, and others. These aspects resulted into more user involvement and hence securing these web applications became incredibly important. This was also due to the fact that the information processed by web applications became very critical to customers, corporations and organizations including countries. There can be very critical information managed through web applications nowadays like financial data, medical records, social securitynumbers, intellectual property and national security data. Webapplications needs to handle this information with utmost care and security.

1.2.1 Cyber-attack

Cyber-attack is any type of offensive manoeuvre employed by individuals or whole organizations that targets computer information systems, infrastructures, computer networks, and/or personal computer devices by various means of malicious acts usually originating from an anonymous source that either steals, alters, or destroys a specified target by hacking into a susceptible system. These can be labelled as either a Cyber campaign, cyber warfare or cyber terrorism in different context. Cyber-attacks can range from installingspyware on a PC to attempts to destroy the infrastructure of entire nations. Cyber-attacks have become increasingly sophisticated and dangerous.

1.2.2 Cyber Warfare and cyber terrorism

Cyber warfare utilizes techniques of defending and attacking information and computer networks that inhabit cyberspace, often through a prolonged Cyber campaign or series of related campaigns. It denies an opponent's ability to do the same, while employing technological instruments of war to attack an opponent's critical computer systems. Cyber terrorism, on the other hand, is "the use of computer network tools to shut down critical national infrastructures (such as energy, transportation, government operations) or to coerce or intimidate a government or civilian population." That means the end result of both cyber

warfare and cyber terrorism is the same, to damage critical infrastructures and computer systems linked together within the confines of cyberspace.

There were two such instances between India and Pakistan that involved cyberspace conflicts, started in 1990s. Earlier cyber-attacks came to known as early as in 1999.Since then, India and Pakistan were engaged in a long-term dispute over Kashmir which moved into cyberspace. Historical accounts indicated that each country's hackers have been repeatedly involved in attacking each other's computing database system. The number of attacks has grown yearly.

1.3 TYPES OF WEB ATTACKS

Firstly, let us have a look into various types of attacks that happen in web arena. Categorizing all web attacks is quite difficult as more and more different ways of attacking gets introduced and evolved. While the security is tightened the attacker also evolve to find more new ways to attack into web. Major types of web attacks are:

i. Spoofing.
ii. Repudiation.
iii. Privacy attacks.
iv. Denial of Service.
v. Privilege escalation.
vi. SQL injection attacks.

1.3.1 Spoofing

1.3.1.1 Email spoofing

Email spoofing (also discussed in next chapter) is the creation of email messages with a forged sender address. It is easy to do because the core protocols do not have any mechanism for authentication. It can be accomplished from within a LAN or from an external environment using Trojan horses. Spam and phishing emails typically use such spoofing to mislead the recipient about the origin of the message.

When an SMTP email is sent, the initial connection provides two pieces of address information:

a. MAIL FROM: - generally presented to the recipient as the Return-path: header but not normally visible to the end user, and by default no checks are done that the sending system is authorized to send on behalf of that address.
b. RCPT TO: - specifies which email address the email is delivered to, is not normally visible to the end user but may be present in the headers as part of the "Received:" header.

Together these are sometimes referred to as the "envelope" addressing, by analogy with a traditional paper envelope.

Once the receiving mail server signals that it accepted these two items, the sending system sends the "DATA" command, and typically sends several header items, including:

From: Joe Q Doe <joeqdoe@example.com> - the address visible to the recipient; but again, by default no checks are done that the sending system is authorized to send on behalf of that address.

Reply-to: Jane Roe <Jane.Roe@example.mil> - similarly not checked

The result is that the email recipient sees the email as having come from the address in the From: header; they may sometimes be able to find the MAIL FROM address; and if they reply to the email it will go to either the address presented in the MAIL FROM: or Reply-to: header - but none of these addresses are typically reliable,so automatedbounce messages may generate backscatter.

Although email spoofing is effective in forging the email address, the IP address of the computer sending the mail can generally be identified from the "Received:" lines in the email header.

1.3.1.2 Website spoofing

Website spoofing is the act of creating a website, as a hoax, with the intention of misleading readers that the website has been created by a different person or organization. Normally, the spoof website will adopt the design of the target website and sometimes has a similar URL. A more sophisticated attack results in an attacker creating a "shadow copy" of the World Wide Web by having all of the victim's traffic go through the attacker's machine, causing the attacker to obtain the victim's sensitive information.

Another technique is to use a 'cloaked' URL.By using domain forwarding, or inserting control characters, the URL can appear to be genuine while concealing the address of the actual website.

The objective may be fraudulent, often associated with phishing or e-mail spoofing, or to criticize or make fun of the person or body whose website the spoofed site purports to represent. Because the purpose is often malicious, "spoof" (an expression whose base meaning is innocent parody) is a poor term for this activity so that more accountable organisations such as government departments and banks tend to avoid it, preferring more explicit descriptors such as "fraudulent" or "phishing".

As an example of the use of this technique to parody an organisation, in November 2006 two spoof websites, www.msfirefox.com and www.msfirefox.net, were produced claiming that Microsoft had bought Firefox and released Microsoft Firefox 2007.

1.3.2 Repudiation

Repudiation makes data or information to appear to be invalid or misleading (Which can even be worse). For example, someone might access your email server and inflammatory information to others under the guise of one of your top managers. This information might prove embarrassing to your companyand possibly do irreparable harm. This type of attack is fairly easy to accomplish because most email systems don't check outbound email for validity. Repudiation attacks like modification attacks usually begin as access attacks.

Non-repudiation refers to a state of affairs where the author of a statement will not be able to successfully challenge the authorship of the statement or validity of an associated contract. The term is often seen in a legal setting wherein the authenticity of a signature is being challenged. In such an instance, the authenticity is being "repudiated".

In a general sense non-repudiation involves associating actions or changes to a unique individual. For a secure area, for example, it may be desirable to implement a key card access system. Non-repudiation would be violated if it were not also a strictly enforced policy to prohibit sharing of the key cards and to immediately report lost or stolen cards. Otherwise determining who performed the action of opening the door cannot be trivially determined. Similarly, for computer accounts, the individual owner of the account must not allow others to use that account, especially, for instance, by giving away their account's password, and a policy should be implemented to enforce this. This prevents the owner of the account from denying actions performed by the account.

1.3.3 Privacy attack

Internet privacy involves the right or mandate of personal privacy concerning the storing, repurposing, provision to third parties, and displaying of information pertaining to oneself via the Internet. Internet privacy is a subset of data privacy. Privacy concerns have been articulated from the beginnings of large scale computer sharing.

Privacy can entail either Personally Identifying Information (PII) or non-PII information such as a site visitor's behaviour on a website. PII refers to any information that can be used to identify an individual. For example, age and physical address alone could identify who an individual is without explicitly disclosing their name, as these two factors are unique enough to typically identify a specific person.

Privacy concerns exist wherever personally identifiable information or other sensitive information is collected and stored – in digital form or otherwise. Improper or non-existent disclosure control can be the root cause for privacy issues. Data privacy issues can arise in response to information from a wide range of sources, such as:

- Healthcare records
- Criminal justice investigations and proceedings
- Financial institutions and transactions
- Biological traits, such as genetic material
- Residence and geographic records
- Ethnicity
- Privacy breach
- Location-based service and geo-location

The challenge in data privacy is to share data while protecting personally identifiable information. The fields of data security and information security design and utilize software, hardware and human resources to address this issue. As the laws and regulations related to

Data Protection are constantly changing, it is important to keep abreast of any changes in the law and continually reassess your compliance with data privacy and security regulations.

Social networking sites try to get users to use their real names, interests, and locations. They believe this makes the social networking experience more realistic, and therefore more engaging for all their users. On the other hand, uploaded photographs or unguarded statements can be identified to an individual, who may regret this exposure. Employers, schools, parents, and other relatives may be influenced by aspects of social networking profiles that the posting individual did not intend for these audiences. On-line bullies may make use of personal information to harass or stalk users. Modern social networking websites allow fine grained control of the privacy settings for each individual posting, but these can be complex and not easy to find or use, especially for beginners.

Photographs and videos posted onto websites have caused particular problems, as they can add a person's face to an on-line profile. With modern and potential facial recognition technology, it may then be possible to relate that face with other, previously anonymous, images, events and scenarios that have been imaged elsewhere. Because of image caching, mirroring and copying, it is difficult to remove an image from the World Wide Web.

1.3.4 Denial of Service

In computing, a denial-of-service (DoS) attack is an attempt to make a machine or network resource unavailable to its intended users, such as to temporarily or indefinitely interrupt or suspend services of a host connected to the Internet. A distributed denial-of-service(DDoS) is where the attack source is more than one–and often thousands–of unique IP addresses.Criminal perpetrators of DoS attacks often target sites or services hosted on high-profile web servers such as banks, credit card payment gateways; but motives of revenge, blackmail or activismcan be behind other attacks.A denial-of-service attack is characterized by an explicit attempt by attackers to prevent legitimate users of a service from using that service. There are two general forms of DoS attacks: those that crash services and those that flood services.

A distributed denial-of-service (DDoS) attack occurs when multiple systems flood the bandwidth or resources of a targeted system, usually one or more web servers.Such an attack is often the result of multiple compromised systems (for example a botnet) flooding the targeted system with traffic. A botnet is a network of zombie computers programmed to receive commands without the owners' knowledge.When a server is overloaded with connections, new connections can no longer be accepted.

1.3.5 Privilege escalation

Privilege escalation is the act of exploiting a bug, design flaw or configuration oversight in an operating system or software application to gain elevated access to resources that are normally protected from an application or user. The result is that an application with more privileges than intended by the application developer or system administrator can perform unauthorized actions.

Most computer systems are designed for use with multiple users. Privileges mean what a user is permitted to do. Common privileges include viewing and editing files, or modifying system files.

Privilege escalation means a user receives privileges they are not entitled to. These privileges can be used to delete files, view private information, or install unwanted programs such as viruses. It usually occurs when a system has a bug that allows security to be bypassed or, alternatively, has flawed design assumptions about how it will be used. Privilege escalation occurs in two forms:

Vertical privilege escalation, also known as privilege elevation, where a lower privilege user or application accesses functions or content reserved for higher privilege users or applications (e.g. Internet Banking users can access site administrative functions or the password for a smartphone can be bypassed).

Horizontal privilege escalation, where a normal user accesses functions or content reserved for other normal users (e.g. Internet Banking User A accesses the Internet bank account of User B).

1.3.6 SQL Injection Attacks

Looking at its wide-spread use in every form of above discussed web attacks SQL injection attack is kept in an altogether separate category of web attacks. SQL injection is a code injection technique, used to attack data-driven applications, in which malicious SQL statements are inserted into an entry field for execution (e.g. to dump the database contents to the attacker). SQL injection must exploit a security vulnerability in an application's software, for example, when user input is either incorrectly filtered for string literal escape characters embedded in SQL statements or user input is not strongly typed and unexpectedly executed. SQL injection is mostly known as an attack vector for websites but can be used to attack any type of SQL database.

1.4 WEB ATTACK FORENSICS

Although there are mechanisms to protect our applications etc. from web attacks but it's quite difficult to find the attacker and book him/her under law. The difficulty in traceability of the hackers/offenders prompts them to do more crimes. The major objective of web forensics is to trace the attacker and in line collect enough evidence that can be presented and accepted in the court of law. The aspects of investigation into web attacks can be viewed in two areas; a) web application forensics and web services forensics.

Although web forensics is a vital necessity the trends are complex and very vast. The increasing adaptability of Service Oriented Architecture (SOAP) in cloud computing scenario has brought lot of scope to the investigation of web services attacks, however, in the current unit we will be focusing more on web application forensics rather than web services. However, we will have little discussion on web services forensics as well here.

1.4.1 Web services forensics

The term "Web services" describes a standardized way of integrating Web-based applications using the XML, SOAP, WSDL and UDDI open standards over an Internet protocol backbone. XML is used to tag the data, SOAP is used to transfer the data, WSDL is used for describing the services available and UDDI lists what services are available.

A Web service is a method of communication between two electronic devices over a network. It is a software function provided at a network address over the Web with the service always on as in the concept of utility computing.Two conceptual elements underlie current web services:

A. Use of XML (eXtensibleMarkup Language), SOAP (Simple Object Access Protocol), and WSDL (Web Service Definition Language) as basic building material;
B. Complex applications built upon long-running transactions that are composed of other web services.

XML format underlies the entire web service architecture and its artefacts. All schemas, definition files, messages transmitted are formed by the means of XML. WSDL, a XML based definition file, defines the interface of a web service in order for the service to be invoked by other services in accordance with the specifications of internal executions. SOAP, a XML based protocol, defines the metadata of the messages to be exchanged between services. WSDL documents define operations; and they are the only mechanisms in order for web services to communicate with each other. Web services use SOAP messages by exchanging them as incoming and outgoing messages through the operations.

There are many attacks on web services, such as WSDL/UDDI scanning, parameter tampering, replays, XML rewriting, man-in-the-middle, eavesdropping, routing detours etc.

As in a document by NIST [csrc.nist.gov/publications/nistir/.../nistir-7559_forensics-web-services.pdf] we need to provide two features into web services forensics:

a) Pairwise evidence generation
b) Comprehensive evidence generation

Pair-wise evidence generation: Collect transactional evidence that occur between pairs of services at service invocation times.

Comprehensive evidence generation: On demand, compose pairs of transactional evidence collected at service invocation times and reveal global views of complex transactional scenarios that occurred during specified periods, and provide them for forensic examiners.

1.4.2 Web Application Forensics

The Major tasks an investigator needs to do while performing web application forensics are:

a) Preliminary analysis: where, we need to focus on evidence collection and protection which are in form of logs. Apart from this the investigator needs to build in confidence by using robust supporting forensic tools. Above all it all depends upon the abilities of the investigator to procure and correlate all data for inferences and conclusion.

b) Standard methodology: methodologies that are standard are easily addressable and heard in the court of law.

1.4.3 Preliminary Analysis

1.4.3.1 Application Forensics Readiness

In this the web application should be well prepared for a forensics investigation. Major activities in it are evidence collection and evidence protection, use of supportive forensics and investigator abilities and more:

i. Evidence collection: A proper Evidence collection is to be done in order to prepare a web application for a forensics investigation.Basically and fore mostly all the logging options of the web application are enabled so as to collect maximum digital evidences. The application logs have to be set according to the case requirement and not be left in default mode which are very basic and might not log important aspects.

ii. Evidence protection: Log files are the main source of digital evidence, hence, proper mechanisms must be incorporated in order to protect these logs files and ensure that these are digitally procured and signed to be presented in court as evidence. This will certainly guarantee the accuracy of the digital evidences provided. Log files can be protected using actions like setting permissions of log files, ensuring out of reach of these log files from the hackers and following checksums to ensure integrity.

iii. Supportive forensics: Mere collection of logs will not help, we must see that these logs are supported by forensics tools evidence gathering. That is, forensic tools can help gather those information which might not be recorded in any application logs. Network or an operating system forensics tools or a third party extra logging facilities can be utilized to achieve this.

iv. Forensics investigator abilities: The forensics investigator must have a sound knowledge and understanding of web application and its architecture etc. better understanding of security aspects and issues pertaining these applications will be required to have a forethought approach in cracking the case.

1.4.3.2 Methodology

Certain prescribed standard methodologies do exist in investigation of web application and these needs to be followed. The cruxes of these standard methodologies are:

a) Protect the web application (could be several servers) during the forensic examination so that their logs etc. can't be modified.
b) Extract all evidence files needed for the forensics investigation:
   - Web servers and application servers logs.
   - Server side scripts which are used by the web application.
   - Web servers and application servers configuration files.
   - All third party software log files.
   - Operating system log files.
c) After collecting the files we need to perform aanalysis of those files to determine the sequence of events and the aspects where security was compromised. One way of carrying out analysis is to divide the log files according to user sessions,by doing this we

will be able to remove distortions and can confine to the culprit's sessions. Fingerprints of a web application security attack needs to be explored. The following are examples of fingerprints and patterns left by web application hacking attempts:

a. Unusual entries in the Logs (GET requests to ASP pages which normally receive POST requests).
b. Script abuse (CMD.exe, Root.exe, Upload. ASP).
c. Excessive attempts from the same IP address.
d. Unusually long processing times (SQL Injection attempt)
e. Files created or modified around the time of the suspected attack. etc.

d) Prepare a report based on the data extracted from the web application logs and other aspects.

1.4.4 Website traffic analysis

Website traffic analysis is produced by grouping and aggregating various data items captured by the web server in the form of log files while the website visitor is browsing the website. Some of the most commonly used website traffic analysis terms are listed below:

URL - A Uniform Resource Locator (URL) uniquely identifies the resource requested by the user's browser.

Hit - Each HTTP request submitted by the browser is counted as one hit. Note that HTTP requests may be submitted for non-existent content, in which case they still will be counted. For example, if one of the five image files referred by the example page mentioned above is missing, the web server will still count six HTTP requests, but in this case, five will be marked as successful (one HTML file and four images) and one as a failed request (the missing image)

Page - A page is a successful HTTP request for a resource that constitutes primary website's content. Pages are usually identified by a file extension (e.g. .html, .php, .asp, etc.) or by a missing extension, in which case the subject of the HTTP request is considered a directory and the default page for this directory is served.

File - Each successful HTTP request is counted as a file.

Visitor - A visitor is the actual person browsing the website. A typical website serves content to anonymous visitors and cannot associate visitors with the actual person browsing the website. Visitor identification may be based on their IP address or an HTTP cookie. The former approach is simple to implement, but results in all visitors browsing the same website from behind a firewall counted as a single visitor. The latter approach requires special configuration of the web server (i.e. to log HTTP cookies) and is more expensive to implement. Note that neither of the approaches identifies the actual person browsing the website and neither provides 100% accuracy in determining that the same visitor has visited the website again.

Visit - A visit is a series of HTTP requests submitted by a visitor with the maximum time between requests not exceeding a certain amount configured by the webmaster, which is typically set at 30 minutes. For example, if a visitor requested page A, then in 10 minutes page B and then in 40 minutes page C, then this visitor has generated two visits, one when pages A and B were requested and another when the page C was requested.

Host - In general, a host is the visitor's machine running the browser. Hosts are often identified by IP addresses or domain names. Those web traffic analysis tools that use IP addresses to identify visitors use the words hosts, domain names and IP addresses interchangeably.

User Agent - User agent is a synonym for a web browser.

1.5 WEB APPLICATION FORENSICS TOOLS

As discussed in the previous section, it is very important to have supportive tools of forensics in order to have better grasp over forensics of web applications. Tools that are useful for web application forensics are Microsoft LogParser, EventLogAnalyzer, Http-analyze, Pyflag, Analog, Open Web Analytics, Mywebalizer, CORE Wisdom, Logjam, Sawmill, and Lire

1.5.1 Logparser

logparser is a flexible command line utility that was initially written by Gabriele Giuseppini, a Microsoft employee, to automate tests for IIS logging. It was intended for use with the Windows operating system, and was included with the IIS 6.0 Resource Kit Tools. The default behaviour of logparser works like a "data processing pipeline", by taking an SQL expression on the command line, and outputting the lines containing matches for the SQL expression.

Microsoft describes Logparser as a powerful, versatile tool that provides universal query access to text-based data such as log files,XML files and CSV files, as well as key data sources on the Windows operating system such as the Event Log, the Registry, the file system, and Active Directory. The results of the input query can be custom-formatted in text based output, or they can be persisted to more specialty targets like SQL, SYSLOG, or a chart.Logparser has been also discussed in previous chapters.

1.5.2 EventLog Analyzer

Event log analysis is used for pattern matching, filtering of event occurrences, and aggregation of event occurrences into composite event occurrences. Commonly,dynamic programming strategies from algorithms are employed to save results of previous analyses for future use, since, for example, the same pattern may be match with the same event occurrences in several consecutive analysis processing.EventLogAnalyzer provides the most cost-effective Security Information and Event Management (SIEM)software on the market. Using this Log Analyzer software, organizations can automate the entire process of managing terabytes of machine generated logs by collecting, analyzing, correlating, searching, reporting, andarchiving from one central location. This event log analyzer software helps to monitor file

integrity, conduct log forensics analysis, monitor privileged users and comply to different compliance regulatory bodies by intelligently analyzing your logs and instantly generating a variety of reports like user activity reports, historical trend reports, and more.

1.5.3 Web log analyzer

Web log analysis software (also called a web log analyzer) is a kind of web analytics software that passes a server log file from a web server, and based on the values contained in the log file, derives indicators about when, how, and by whom a web server is visited. Usually reports are generated from the log files immediately, but the log files can alternatively be passed for a database and reports generated on demand.

1.5.4 Open Web Analytics[8]

Open Web Analytics (OWA) is open source web analytics software created by Peter Adams. OWA is written in PHP and uses aMySQL database, which makes it compatible for running with an AMP solution stack on various web servers.OWA is comparable to Google Analytics, though OWA is server software anyone can install and run on their own host, while Google Analytics is asoftware service offered by Google.OWA supports tracking with WordPress and MediaWiki, two popular web site frameworks.This application helps you keep track of and observe the influx of views on your website. The program also tracks your competitors and their company's growth compared to yours.

1.5.5 Webalizer

The **Webalizer** is a GPL application that generates web pages of analysis, from access and usage logs, i.e. it is web log analysis software. It is one of the most commonly used web server administration tools. It was initiated by Bradford L. Barrett in 1997. Statistics commonly reported by Webalizer include hits, visits, referrers, the visitors' countries, and the amount of data downloaded. These statistics can be viewed graphically and presented by different time frames, such as by day, hour, or month.

1.6 SUMMARY

1. There can be very critical information managed through web applications nowadays which needs to handle this information with utmost care and security.
2. Cyber-attacks have become increasingly sophisticated and dangerous. Cyber-attacks have evolved into Cyber Warfare and cyber terrorism.
3. Various forms of cyber-attacks are, Spoofing, Repudiation, Privacy attacks, Denial of Service, Privilege escalation, SQL injection attacks.
4. The aspects of investigation into web attacks can be viewed in two areas; web application forensics and web services forensics.
5. Web services forensics involves pair-wise and comprehensive evidence generation. Web application forensics involves preliminary analysis and standard methodologies.
6. Tools that are useful for web application forensics are Microsoft LogParser, EventLogAnalyzer, Http-analyze, Pyflag, Analog, Open Web Analytics etc.

1.7 CHECK YOUR PROGRESS

1. Fill in the blanks.

a) _____ is the creation of email messages with a forged sender address.
b) _____and_____ emails are typical means of spoofing.
c) IP address of the computer sending the mail can generally be identified from the _____ in the email header.
d) _____ refers to a state of affairs where the author of a statement will not be able to successfully challenge the authorship of the statement or validity of an associated contract.
e) _____ is where a lower privilege user or application accesses functions or content reserved for higher privilege users or applications.
f) Application Forensics Readiness involves _____, _____, _____ and _____.

2. State True or False

a) Repudiation does not make data or information to appear to be invalid or misleading.
b) User agent is a synonym for a web browser.
c) Open Web Analytics (OWA) is open source web analytics software created by Peter Adams
d) Open Web Analytics is written in PHP and uses a MySQL database.
e) A distributed denial-of-service (DDoS) attack occurs when multiple systems flood the bandwidth or resources of a targeted system, usually one or more web servers

1.8 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) Email spoofing.
b) Spam and phishing.
c) Received sub-head.
d) Non-repudiation.
e) Vertical privilege escalation.
f) evidence collection, evidence protection, use of supportive forensics and investigator abilities.

2. State True or False

a) False.
b) True.
c) True.
d) True.
e) True.

1.9 FURTHER READINGS

1. Computer Forensics: Investigating Network Intrusions and Cyber Crime, EC-Council, Cengage learning 2010.
2. KeyunRuan, Cybercrime and Cloud Forensics: Applications for Investigation Processes, IGI Global, 2013.
3. Gutiérrez, Carlos A., Web Services Security Development and Architecture: Theoretical and Practical issues, IGI Global, 2010.

4. Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
5. Amor Lazzez, ThabetSlimani Forensics Investigation of Web Application Security Attacks I. J. Computer Network and Information Security, 2015, 3, 10-17

1.10 MODEL QUESTIONS

1. What do you mean by "shadow copy" of the World Wide Web and 'cloaked' URL? Where does it take place? Describe in detail.
2. Describe the Major tasks an investigator needs to do while performing web application forensics?
3. Describe the major types of web attacks in brief.
4. What do you mean by Application Forensics Readiness?
5. Describe any 3 web application forensic tools.

References, Article Source & Contributors

[1] Cyber-attack - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Cyber-attack
[2] Denial-of-service attack - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Denial-of-service_attack
[3] Email spoofing - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Email_spoofing
[4] Event monitoring - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Event_monitoring
[5] Forensics Web Services - NIST Computer Security, csrc.nist.gov/publications/nistir/.../nistir-7559_forensics-web-services.pdf
[6] Open Web Analytics - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Open_Web_Analytics
[7] Privacy issues of social networking sites - Wikipedia, https://en.wikipedia.org/wiki/Privacy_issues_of_social_networking_sites
[8] Privilege escalation - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Privilege_escalation
[9] SQL injection - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/SQL_injection
[10] Types of Attacks, “Ethical hacking Tips”, Go4Expert, www.go4expert.com
[11] Web log analysis software - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Web_log_analysis_software
[12] Web service - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Web_service
[13] Webalizer - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Webalizer

UNIT II: INVESTIGATING EMAIL ATTACKS

2.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand emailing and email services.
- Know the structure of email and correlate it to extract forensic information.
- Know and categorize email attacks and crimes.
- Know and familiarise with few email forensic tools.

2.2 INTRODUCTION

Email and web based email or web mail has a strong tendency to spread rapidly to a larger extents. Email is nowadays considered as an evidence and has been a reason for putting good amount of people(criminals) behind bars. Laws nowadays gives importance to emails and reviews them with lot of attention.

2.2.1 Email Structure

An Internet email message consists of three components, the message envelope, the message header, and the message body. The message header contains control information, including, minimally, an originator's email address and one or more recipient addresses. Usually descriptive information is also added, such as a subject header field and a message submission date/time stamp.

Originally an ASCII text-only communications medium, Internet email was extended by Multipurpose Internet Mail Extensions (MIME) to carry text in other character sets and multi-media content attachments. International email, with internationalized email addresses using UTF-8, has been standardized, but not yet widely adopted.

Internet email messages consist of two major sections, the message header and the message body. The header is structured into fields such as From, To, CC, Subject, Date, and other information about the email. The body contains the message, as unstructured text, sometimes containing a signature block at the end. The header is separated from the body by a blank line.

2.2.2 Types of email services

2.2.2.1 Web-based email

Many email providers have a web-based email client (e.g. AOL Mail, Gmail, Outlook.com and Yahoo! Mail). This allows users to log into the email account by using any compatible web browser to send and receive their email. Mail is typically not downloaded to the client, so can't be read without a current Internet connection.

2.2.2.2 POP3 email services

The Post Office Protocol 3 (POP3) is a mail access protocol used by a client application to read messages from the mail server. Received messages are often deleted from theserver. POP supports simple download-and-delete requirements for access to remote mailboxes IMAP email servers

2.2.2.3 The Internet Message Access Protocol (IMAP)

It provides features to manage a mailbox from multiple devices. Small portable devices like smartphones are increasingly used to check email while travelling, and to make brief replies, larger devices with better keyboard access being used to reply at greater length. IMAP shows the headers of messages, the sender and the subject and the device needs to request to download specific messages. Usually mail is left in folders in the mail server.

2.2.2.4 MAPI email servers

Messaging Application Programming Interface (MAPI) is a messaging architecture and an API based on the Component Object Model (COM) for Microsoft Windows.

2.2.3 Simple Mail Transfer Protocol

Simple Mail Transfer Protocol (SMTP) is an Internet standard for electronic mail (email) transmission. SMTP by default uses TCP port 25. The protocol for mail submission is the same, but uses port 587. SMTP connections secured bySSL, known as SMTPS, default to port 465 (nonstandard, but sometimes used for legacy reasons).

Although electronic mail servers and other mail transfer agents use SMTP to send and receive mail messages, user-level client mail applications typically use SMTP only for sending messages to a mail server for relaying. For receiving messages, client applications usually use either POP3 or IMAP. Although proprietary systems (such as Microsoft Exchange and IBM Notes) and webmail systems (such as Outlook.com, Gmail andYahoo! Mail) use their own non-standard protocols to access mail box accounts on their own mail servers, all use SMTP when sending or receiving email from outside their own systems.

Figure 33 gives a typical sequence of events that takes place when sender Alice transmits a message using a Mail User Agent (MUA) addressed to the email address of the recipient.

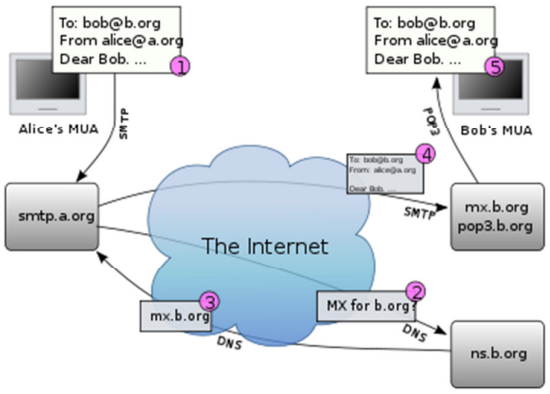

Figure 33: SMTP Scenario (from Wikipedia)

- The MUA formats the message in email format and uses the submission protocol, a profile of the Simple Mail Transfer Protocol (SMTP), to send the message to the local mail submission agent (MSA), in this casesmtp.a.org.

- The MSA determines the destination address provided in the SMTP protocol (not from the message header), in this case bob@b.org. The part before the @ sign is the local part of the address, often theusername of the recipient, and the part after the @ sign is a domain name. The MSA resolves a domain name to determine the fully qualified domain name of the mail server in the Domain Name System (DNS).
- The DNS server for the domain b.org (ns.b.org) responds with any MX records listing the mail exchange servers for that domain, in this case mx.b.org, a Message Transfer Agent (MTA) server run by the recipient's ISP.
- smtp.a.org sends the message to mx.b.org using SMTP. This server may need to forward the message to other MTAs before the message reaches the final message delivery agent (MDA).
- The MDA delivers it to the mailbox of user bob.
- Bob's MUA picks up the message using either the Post Office Protocol (POP3) or the Internet Message Access Protocol (IMAP).

Alice or Bob may use a client connected to a corporate email system, such *as IBM Lotus Notes* or *Microsoft Exchange*. These systems often have their own internal email format and their clients typically communicate with the email server using a vendor-specific, proprietary protocol. The server sends or receives email via the Internet through the product's Internet mail gateway which also does any necessary reformatting. If Alice and Bob work for the same company, the entire transaction may happen completely within a single corporate email system.

- Alice may not have a MUA on her computer but instead may connect to a webmail service.
- Alice's computer may run its own MTA, so avoiding the transfer at step 1.
- Bob may pick up his email in many ways, for example logging into mx.b.org and reading it directly, or by using a webmail service.
- Domains usually have several mail exchange servers so that they can continue to accept mail even if the primary is not available.

Many MTAs used to accept messages for any recipient on the Internet and do their best to deliver them. Such MTAs are called open mail relays. This was very important in the early days of the Internet when network connections were unreliable. However, this mechanism proved to be exploitable by originators of unwanted bulk email and as a consequence open mail relays have become rare, and many MTAs do not accept messages from open mail relays.

2.3 EMAIL ATTACKS AND CRIMES

Email crimes or attacks can be a direct one where users can use them to harass or intimidate a receiver. There exist lots of crimes which are perpetrated directly using emails. Also email attacks can be indirect where emailing is used as one of the tool to capture sensitive information and perform malpractices or induce malwares into the client system. Let us look into few email attacks or crimes.

a. Flaming
b. Email spoofing

c. Email bombing
d. Email hacking
e. Spams
f. Email frauds
g. Email phishing

2.3.1 Flaming

Flamming occurs when a person sends a message with angry or antagonistic content. The term is derived from the use of the word Incendiary to describe particularly heated email discussions. Flaming is assumed to be more common today because of the ease and impersonality of email communications: confrontations in person or via telephone require direct interaction, where social norms encourage civility, whereas typing a message to another person is an indirect interaction, so civility may be forgotten.

2.3.2 Email spoofing

It occurs when the email message header is designed to make the message appear to come from a known or trusted source. Email spam and phishing methods typically use spoofing to mislead the recipient about the true message origin.

2.3.3 Email bombing

It is the intentional sending of large volumes of messages to a target address. The overloading of the target email address can render it unusable and can even cause the mail server to crash.

2.3.4 Email hacking

It is illicit access to an email account or email correspondence.

2.3.5 Spams

Attackers often send massive email broadcasts with a hidden or misleading incoming IP address and email address.Some users may open the spam, read it, and possibly be tempted by whatever wares or schemes are offered.

2.3.6 Phishing

This type of attacks uses email messages from legitimate businesses that the user may be associated with. Although the messages look authentic with all the corporate logos and similar format as the official emails, they ask for verification of personal information such as the account number, password, and date of birth. 20% of unsuspecting victims respond to them, which may result in stolen accounts, financial loss and identity theft.

2.3.7 Email fraud

It is the intentional deception made for personal gain or to damage another individual through email. Almost as soon as email became widely used, it began to be used as a means to defraud people. Email fraud can take the form of a "con game" or *scam*. Confidence tricks tend to exploit the inherent greed and dishonesty of their victims. The prospect of a 'bargain' or 'something for nothing' can be very tempting. Email fraud, as with other 'bunco schemes' usually targets naive individuals who put their confidence in get-rich-quick schemes such as

'too good to be true' investments or offers to sell popular items at 'impossibly low' prices. Many people have lost their life savings due to fraud.

2.3.8 Phishing emails

It may contain links to websites that are infected with malware.Phishing is typically carried out by email spoofingor instant-messaging,and it often directs users to enter details at a fake website whose look and feel are almost identical to the legitimate one. Phishing is an example of social engineering techniques used to deceive users, and exploits the poor usability of current web security technologies.

2.4 PRIVACY IN EMAILS

2.4.1 Email privacy

It is the broad topic dealing with issues of unauthorized access and inspection of electronic mail. This unauthorized access can happen while an email is in transit, as well as when it is stored on email servers or on a user computer. In countries with a constitutional guarantee of the secrecy of correspondence, whether email can be equated with letters and get legal protection from all forms of eavesdropping comes under question because of the very nature of email. This is especially important as more and more communication occurs via email compared to postal mail.

Email has to go through potentially untrusted intermediate computers (email servers, ISPs) before reaching its destination, and there is no way to tell if it was accessed by an unauthorized entity. This is different from a letter sealed in an envelope, where by close inspection of the envelope, it might be possible to tell if someone opened it. In that sense, an email is much like a postcard whose contents are visible to everyone who handles it.

There are certain technological workarounds that make unauthorized access to email hard, if not impossible. However, since email messages frequently cross nation boundaries, and different countries have different rules and regulations governing who can access an email, email privacy are a complicated issue.

A significant fraction of email communication is still unencrypted. In general, encryption provides protection against malicious entities. However, a court order might force the responsible parties to hand over decryption keys;

- Email privacy, without some security precautions, can be compromised because:
- Email messages are generally not encrypted.
- Email messages have to go through intermediate computers before reaching their destination, meaning it is relatively easy for others to intercept and read messages.
- Many Internet Service Providers (ISP) store copies of email messages on their mail servers before they are delivered. The backups of these can remain for up to several months on their server, despite deletion from the mailbox.
- The "Received:"-fields and other information in the email can often identify the sender, preventing anonymous communication.

2.4.2 Email tracking

It is a method for monitoring the email delivery to intended recipient. Most tracking technologies use some form of digitally time-stamped record to reveal the exact time and date that an email was received or opened, as well the IP address of the recipient.

Email tracking is useful when the sender wants to know if the intended recipient actually received the email, or if they clicked the links. However, due to the nature of the technology, email tracking cannot be considered an absolutely accurate indicator that a message was opened or read by the recipient.

2.5 EMAIL FORENSICS

2.5.1 Forensically important email parts

Basically emails information which will be interesting to the investigators are:

a) Email header
b) Body of Emails
c) The information hidden in the email packets
d) Attachments

The message header must include at least the following fields:

- *From*: The email address, and optionally the name of the author(s). In many email clients not changeable except through changing account settings.
- *Date*: The local time and date when the message was written. Like the *From:* field, many email clients fill this in automatically when sending. The recipient's client may then display the time in the format and time zone local to him/her.

The message header should include at least the following fields:

- *Message-ID*: Also an automatically generated field; used to prevent multiple deliveries and for reference in In-Reply-To: (see below).
- *In-Reply-To*: Message-ID of the message that this is a reply to. Used to link related messages together. This field only applies for reply messages.

RFC 3864 describes registration procedures for message header fields at the IANA; it provides for permanent and provisional message header field names, including also fields defined for MIME, netnews, and http, and referencing relevant RFCs. Common header fields for email include:

- *To*: The email address(es), and optionally name(s) of the message's recipient(s). Indicates primary recipients (multiple allowed), for secondary recipients see Cc: and Bcc: below.
- *Subject*: A brief summary of the topic of the message. Certain abbreviations are commonly used in the subject, including "RE:" and "FW:".
- *Bcc*: Blind carbon copy; addresses added to the SMTP delivery list but not (usually) listed in the message data, remaining invisible to other recipients.

- *Cc*: Carbon copy; Many email clients will mark email in one's inbox differently depending on whether they are in the To: or Cc: list.
- Content-Type: Information about how the message is to be displayed, usually a MIME type.
- *Precedence*: commonly with values "bulk", "junk", or "list"; used to indicate that automated "vacation" or "out of office" responses should not be returned for this mail, e.g. to prevent vacation notices from being sent to all other subscribers of a mailing list. Sendmail uses this header to affect prioritization of queued email, with "Precedence: special-delivery" messages delivered sooner. With modern high-bandwidth networks delivery priority is less of an issue than it once was. Microsoft Exchange respects a fine-grained automatic response suppression mechanism, the X-Auto-Response-Suppress header.
- *References*: Message-ID of the message that this is a reply to, and the message-id of the message the previous reply was a reply to, etc.
- *Reply-To*: Address that should be used to reply to the message.
- *Sender*: Address of the actual sender acting on behalf of the author listed in the From: field (secretary, list manager, etc.).
- *Archived-At*: A direct link to the archived form of an individual email message.

SMTP defines the *trace information* of a message, which is also saved in the header using the following two fields:

- *Received*: when an SMTP server accepts a message it inserts this trace record at the top of the header (last to first).
- *Return-Path*: when the delivery SMTP server makes the *final delivery* of a message, it inserts this field at the top of the header.

Other header fields that are added on top of the header by the receiving server may be called *trace fields*, in a broader sense.

- *Authentication-Results*: when a server carries out authentication checks, it can save the results in this field for consumption by downstream agents.
- *Received-SPF*: stores results of Sender Policy Framework (**SPF**) checks in more detail than Authentication-Results.
- *Auto-Submitted*: is used to mark automatically generated messages.
- *VBR-Info*: claims VBR whitelisting. Vouch by Reference (VBR) is a protocol for adding third-party certification to email.

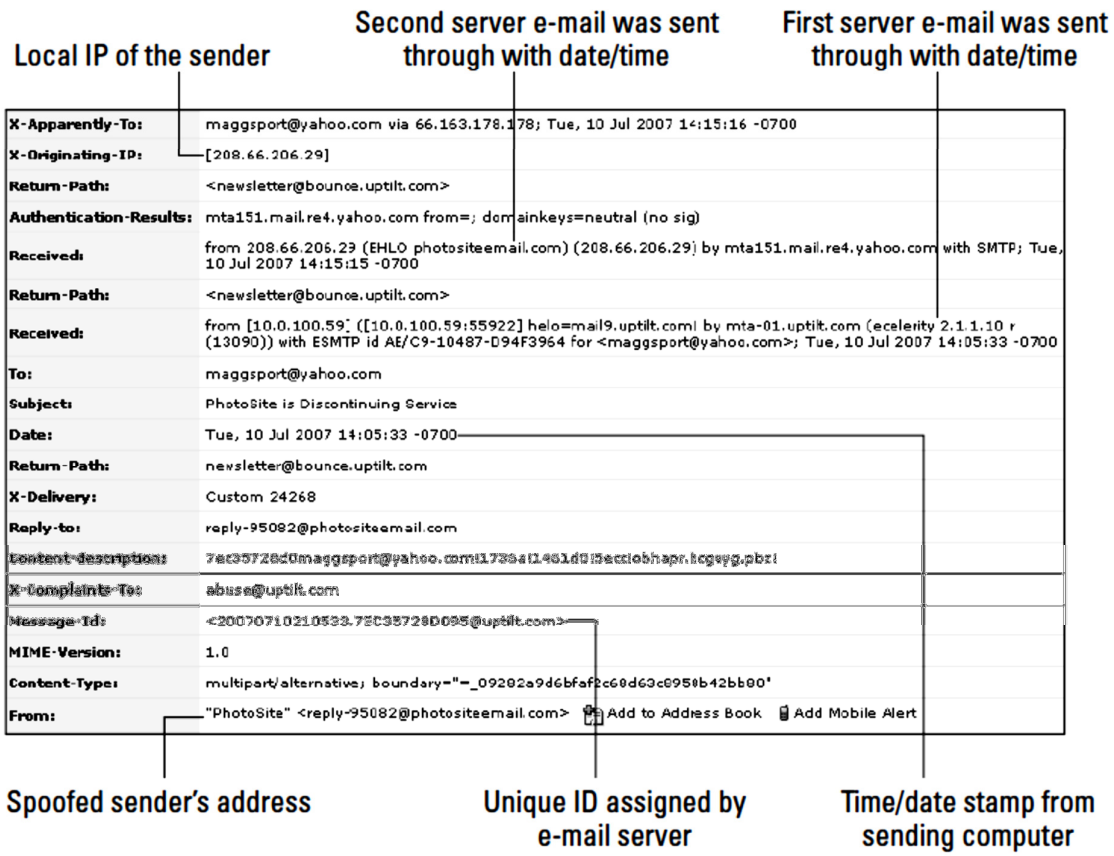

Figure 34: Tracing spoofed sender.

The trace information of an email can provide lots of clues to the investigators.

The email packets can be captured using packet sniffer software. The email packets can be read very easily unless the user is having email encryption. The encrypted emails are read using the password cracking methodologies as discussed in earlier chapters. The trace of an email, headers and even sometimes the body of the email can be used detect a spoof attack as shown in *Figure 34*.

2.5.2 Email forensics investigation

Email forensics involves capturing, securing and analysing and reporting the email evidences. E-mail forensics aims to study the source and contentsof e-mail messages for evidence, this included identification of the actual sender, recipient, date and time when it wassent, etc. Email Forensic analysis aims at discovering the history ofa message and confirming identity of all involved entities. Apart from message analysis, e-mailforensic also involves investigation of clients or server computers suspectedof being used or misused to carry out e-mail forgery. It might involve inspection of *Internetfavorites*, *Cookies, History, Typed URL's, Temporary Internet Files, Auto-completionEntries, Bookmarks, Contacts, Preferences, Cache*, etc. Several OpenSource software tools are available which helps to perform e-mail header analysisto collect evidence of e-mail fraud.

2.5.3 Analyzing an email

A sample header set of an e-mail message sent by *abc@xyz.com* pretending to be *alice@alice.com* and sent to *bob@bob.com* is shown in figure 3.

1 **X-Apparently-To:** bob@bob.com via a4.b4.c4.d4; Tue, 30 Nov 2010 07:36:34 -0800
2 **Return-Path:** < alice@alice.com >
3 **Received-SPF:** none (mta1294.mail.mud.bob.com: domain of alice@alice.com does not designate permitted sender hosts)
4 **X-Spam-Ratio:** 3.2
5 **X-Originating-IP:** [a2.b2.c2.d2]
6 **X-Sieve:** CMU Sieve 2.3
7 **X-Spam-Charsets:** Plain='utf-8' html='utf-8'
8 **X-Resolved-To:** bob@bob.com
9 **X-Delivered-To:** bob@bob.com
10 **X-Mail-From:** alice@alice.com
11 **Authentication-Results:** mta1294.mail.mud.bob.com from=alice.com; domainkeys=neutral (no sig); from=alice.com; dkim=neutral (no sig)
12 **Received:** from 127.0.0.1 (EHLO mailbox-us-s-7b.xyz.com) (a2.b2.c2.d2) by mta1294.mail.mud.bob.com with Journal of Digital Forensics, Security and Law, Vol. 6(2)56 SMTP; Tue, 30 Nov 2010 07:36:34 -0800
13 **Received:** from MTBLAPTOP (unknown [a1.b1.c1.d1]) (Authenticated sender: abc@xyz.com) by mailbox-us-s-7b.xyz.com (Postfix) with ESMTPA id 8F0AE139002E for <bob@bob.com>; Tue, 30 Nov 2010 15:36:23 +0000 (GMT)
14 **From:** "Allice" <Alice@a.com>
15 **Subject:** A Sample Mail Message
16 **To:** "Bob Jones" <bob@bob.com>
17 **Content-Type:** multipart/alternative; charset="utf-8"; boundary="KnRI8MgwQQWMSCW6Q5=_ HgI2hw Adah5NLY"
18 **MIME-Version:** 1.0
19 **Content-Transfer-Encoding:** 8bit
20 **Content-Length:** 511
21 **Reply-To:** "Smith" <smith@smith.com>
22 **Organization:** Alices Organization
23 **Date:** Tue, 28 Nov 2010 21:06:22 +0530
24 **Return-Receipt-To:** smith@smith.com
25 **Disposition-Notification-To:** jones@jones.com
26 **Message-Id:** <20101130153623. 8F0AE139002E@mailbox-us-s-7b.xyz.com>

Figure 35: Elaborate email header of a spoofed email. (adapted from: [6])

The Header *X-Apparently-To* shown in Figure 35 is relevant when mail has been sentas a BCC or to recipients of some mailing list. This field in most of the casescontain the address as in Tofield. But if mail has been sent to a BCC recipient ora mailing list, *X-Apparently-To* is different from *TO* field. Some may show *TO*while others may not show it. Thus *X-Apparently-To* always shows the e-mailaddress of recipient regardless of whether mail has been sent using *TO*, BCC, CCaddresses or by the use of some mailing list.

The *Return-Path* header is the e-mail address of the mailbox specified by thesender in the *MailFrom*command. This address can also be spoofed, if noauthentication mechanism is in place at the sending server it is not possible to determinegenuineness of Return-Path header through header analysis alone.The *Received-SPF* specifies thatthe mail has come from a domain which either does not have a SPF record or isnot yet a designated permitted sender.

If there are some spam filtering software of the receiving serveror MUA the spam score is contained in *X-Spam-Ratio* field. If this value for the e-mail under study ratio exceeds certain pre-defined threshold, emailwill be classified as spam.

X-Originating-IP specified the IP address of the last MTA of the sending SMTPServer, which has delivered the e-mail to the server of *bob@bob.com*. In thesample e-mail it is *[a2.b2.c2.d2]* as shown in item 5. This address is alsocontained in the *Received* header field.*X-Sieve* header specifies the name and version of message filtering system. Thispertains to the scripting language used to specify conditions for message filteringand handling. In the sample e-mail the name of the message filtering software is*CMU Sieve* and its version is *2.3.X-Spam-Charsets* header specifies the character set used for filtering themessages. The value for this field in sample e-mail at item 7 indicates that 8-bitUnicode Transformation Format (*UTF*) has been used by bob's server. UTF is avariable length character set having a special property of being backwardcompatibleto ASCII.*X-Resolved-To* address is the e-mail address of the mailbox to which the mail hasbeen delivered by MDA of bob's server. In most cases, it is the same as XDelivered-To field. *X-Delivered-To* is the address of the mailbox to which themail has been delivered by MDA of bob's server. In the sample e-mail both *XResolved-To and X-Delivered-To* addresses are *bob@bob.com* as in item 8 and 9.*X-Mail-From* header specifies the e-mail address of the mailbox specified by thesender in the *MailFrom*command which in the sample e-mail is alice@alice.com.The *Authentication-Results* header in item 11 indicates that*mta1294.mail.mud.bob.com* received mail from *alice.com* domain which neitherhas DomainKeys signature nor DKIM signature.item 12 is the second *Received* header field containing the trace informationindicating *127.0.0.1* as the IP address of the machine that send the message. Thismachine is actually named *mailbox-us-s-7b.xyz.com* and has IP address*a2.b2.c2.d2*. It has used *EHLO* SMTP command to send the mail. The mail wasreceived by *mta1294.mail.mud.bob.com* using *SMTP*. The message has beenreceived on *Tue, 30 Nov 2010* date at *07:36:34* time. The clock is 8 hrs behindGreenwich Mean Time. Item 13 is the first *Received* header field representing the trace informationindicating *MTBLAPTOP* as the names of the machine that send the message. Thismachine is not known to the receiver but has an IP address *a1.b1.c1.d1* and*abc@xyz.com* is the owner of the mailbox who has sent the message. TheMTA must follow some authentication mechanism to identify its mailbox usersotherwise it is not possible to include authenticated sender's mailbox address withthe *Received* field. The message has been received by *mailbox-us-s-7b.xyz.com*using *ESMTPA* protocol which has been running a program called *Postfix*. Themessage is for *bob@bob.com* and has an ID of *8F0AE139002E*. The message hasbeen received on *Tue, 30 Nov 2010 at 15:36:23*. The clock is set according toGreenwich Mean Time.

The *From*, *Subject* and *To*lines respectively are the e-mail address of the author,subject of the message, and the e-mail address of the intended recipient. *Subject*and *To*are specified by the sender, and the *From* address is taken by the systemfrom the current logged in user. However, *From*header can very easily bespoofed as has been dome in this sample e-mail. The items 14, 15 and 16 inthe sample e-mail show the values of these three fields. The *From*address hasbeen spoofed to carry an address *Alice@a.com* with a user friendly name *Alice*.Content-Type, MIME-Version, Content-Transfer-Encoding and Content-length in items 17, 18, 19 and 20 are the MIME headers describing the type of MIMEcontent, transfer encoding, its version and length so that the MUA's can performproper decoding to render the message successfully on client.This is the address, sender of this e-mail wants recipient to use for sending replyin response to this e-mail. Normally, this is used by the senders to send replies.Carefully crafted

sender spoofing combined with fake *Reply-To* e-mail addresscan lead to serious information leaks. The *Reply-To* address *"Smith"smith@smith.com* in item 21 is an arbitrary address that may belong to some userwho may not be related to the sender in any way.

Organization header field indicates that the organization of claimed sender is*Alices Organization*. *Organization* header field is an information fieldrepresenting the organization of a sender. It can be misused by the spammer togive a false impression about a sender as has been done in this e-mail.

Date header indicates that the e-mail was composed and submitted for delivery on*Tue, 28 Nov 2010 21:06:22 +0530*, which is not in conformity with the date inthe *Received* field of Para 23.*Return-Receipt-To* field indicates the e-mail address, MSA, MTA and MDA mustuse for sending delivery notifications such as successful or failure notifications.The address mentioned for this field in item 24 is again an arbitrary address thatmay belong to some user who may not be related to the sender in any way.*Disposition-Notification-To* field indicates an e-mail address, MUA must usewhen submitting a message indicating that the message has been displayed. This address specified in item 25 is also an arbitrary address that does belong to someuser who may not be related to the sender in any way.item 26 contains the *Message-Id* of the message which is

20101130153623.8F0AE139002E@mailbox-us-s-7b.xyz.com. Generally, adomain name is appended with a unique number by the sending server to form the*Message-Id*. In the above sample e-mail message, several fields have been spoofed which canbe detected easily because the first *Received* field shows the address ofauthenticated sender which is different from the sender of the message. However,address of authenticated sender may not be always included with theauthentication results (in case no authentication mechanism is adhered to orannomizers strip this line). Further, date is also inconsistent as can be noted fromthe comparison of timestamp in *Received* headers and the date field. Some headerfields with context to authentication and above analysed e-mail message arediscussed further hereby:

SPF mechanisms can be used to describe the set of hosts which are designatedoutbound mailers for the domain. The test besides success or failure may alsoresult into *softfail, neutral, none, permerror*or *temperror*. For example, asuccessful *Received-SPF* entry could be as follows:

Received-SPF: pass (mta1104.mail.mud.xyz.com: domain of abc@xyz.com designates a2.b2.c2.d2 as permitted sender)Here, the *mta1104.mail.mud.xyz.com MTA* notifies its recipient through*Received-SPF* that domain of abc@xyz.com i.e. *xyz.com* which has an *IP*address *a2.b2.c2.d2* is a permitted sender designated by Sender Policy Framework. In case, the domain *alice.com* had usedDomainKeys and DKIM complaint andhad passed these tests, it could have been as follows:

> ***Authentication-Results:*** mta1294.mail.mud.bob.com from=alice.com;
> domainkeys=pass (ok); from=a.com; dkim=pass (ok)
> In this case, it could have included DKIM-Signature and/orDomainKey-Signature fields as follows:
> ***DKIM-Signature:*** v=1; a=rsa-sha1; c=simple; d=alice.com;
> h=from:to:subject:date:message-id:content-type q=dns/txt; s=s512;
> bh=XX……=; b=XXX……==;

This is the DKIM Signature signed with SHA1 algorithm. DKIM uses the emailheaders and body to generate a signature. If the headers are rewritten or text is appended to the message body after it has been signed, the DKIM verificationfails. DKIM is backward compatible with the DomainKeys system. When an emailmessage is signed with DKIM, it will include a number of "tags" whosevalues contain authenticating data for the message being sent. In the example email header in figure 3, the tags used are:

v= This tag defines the version of this specification that applies to thesignature record.

a= The algorithm used to generate the signature (plain-text;REQUIRED). It supports "rsa-sha1" and "rsa-sha256", Signersusually signs using "rsa-sha256".

c= It is the canonicalization algorithm 1.e. the method by which theheaders and content are prepared for presentation to the signingalgorithm.

d= It is the domain name of the signing domain.

h= It is a colon-separated list of header field names that identify theheader fields presented to the signing algorithm.

q= It specifies the query method used to retrieve the public key whichby default is dns.

s= It is the selector used in the public key.

bh= The signature data or public key, encoded as a Base64 string.

The example of DomainKeys signature is given below. DomainKeys signaturehas been signed with SHA1 algorithm.

DomainKeys-Signature: a= rsa-sha1; q=dns; c=simple; s=s512;
d=alice.com; b=XXX……==;

When an e-mail message is signed with DomainKeys, it will include a number of"tags" whose values contain authenticating data for the message being sent. In theexample above, the tags used are:

a= It is the encryption algorithm used to generate the signature which bydefault is "rsa-sha1".

q= It specifies the query method used to retrieve the public key which bydefault is dns.

c= It is the canonicalization algorithm 1.e. the method by which theheaders and content are prepared for presentation to the signingalgorithm.

s= It is the selector used in the public key.

d= It is the domain name of the signing domain.

b= The signature data or public key, encoded as a Base64 string.

Date header represents the date e-mail was composed and submitted for delivery.However, this filed can also be spoofed as has been done in thissample e-mail message. It can be easily noticed by comparing its value in item 23with the dates in the *Received* header fields.

Message-Id is the message Identification attached to the e-mail message. Every e mail has a unique message ID that helps the administrators to locate the e-mail inserver log. Usually every sending server uses its own custom algorithm togenerate this unique number and append domain name to this to make it uniqueon the internet. This ID can also help to identify the domain of the sender but itcan also be forged to confuse the investigators.

The first *Received* header field representing the trace information contains the IPaddress of the machine used to send the e-mail message. On tracking this IPaddress several cases as explained below are possible:

i. The IP address in the *Received* header field maps to directconnection having a static IP address. In this case, this address isthe address of the sender's computer. However, if the IP address isdynamic then the logs of the proxy or SMTP server need to beobtained for continuing the e-mail tracking.

ii. The IP address contained in the *Received* header corresponds tosome proxy server. In this case, proxy server's log must be obtainedto track the sender. Open proxy server may raise some issues for theinvestigators because they do not maintain a strict log of activities.In case SSL is used to log on to *HTTP* based e-mail server, proxycannot be an issue because IP address of the client shall berecorded. Corporate proxy servers may not be strictly timesynchronized as they may be using Network Time Protocol (*NTP)*and thus may impede the investigation. *ISP* proxy servers usuallymaintain a strict and time synchronized log (using*STIME* protocol)and have a clear devised policy to cooperate with the investigators.

iii. The tracked IP address maps to some tunnelling server. In this case,tracking source of e-mail will be difficult because tunnelling maybe done in different ways and some are not logged.

iv. The IP address in the *Received* header field maps to SMTP server.In this case, the SMTP server log must be obtained. IP address maymap to SMTP server belonging to ISP, or some corporate or anopen relay. In all cases, logs stored must be obtained. If the logs arestrictly time synchronized, then the sender can be tracked easily.ISP and corporate SMTP servers can provide further details aboutthe particular user such as his contact details and credit card number.

v. The IP address contained in the *Received* field resolves to Annomizers or re-mailers. In this case, investigators must obtainlogs and original e-mail message from the anonymous SMTP orHTTP servers. Further, in case the anonymity is a paid service, useraccount details must also be obtained. It is also possible to add one or more false *Received* headers in the data field ofthe message with an intention to freeze the investigation. Investigators must paycareful attention to all fields of the *Received* headers with respect to each otherespecially in terms of delivery methods and date & time. If the delivery methodsvary or the time & date differ considerably, then false headers can be easilyidentified. Otherwise, the investigation shall have to investigate all IP addressesand request logs from all servers. It may be very difficult to track a sender fromthe IP address if the sender has tampered IP address at packet level. Once the source of the e-mail message under investigation has beendetermined or someone is strongly suspected for being the source, his or hercomputer, e-mail client software, web browser, etc. are investigated for traces ofevidence.

2.5.4 Instant Messages

Instant Messages (IM) (as mostly referred as chats) has been becoming very popular among users. Emails are mostly attached to inboxes whereas the IMs are based on text cells or forms. Texting on mobile devices has become very popular nowadays with apps like Whatsapp.

IMs too are very important to forensic examiners because nowadays companies are using this form of communicationfor real-time customer service and internal business communication.On the peopleperspective, IMs are used to chat about everything from recipes

to personal attributes or opinions. Chats are relayed by way of a server. Same goes for IMs too. IM software are structurally same as e-mail systems the only difference is that IMs are done in real time.
at real-timesits necessary to logthe data (communication) as it is being typed. Recovering chat sessions is a matter of chance because the caching abilities of the computer is the elementthat is required to re-create the chat sessions. Some IM software logs conversations, but generally people don't activate the logs. IMsare migrating to mobile devices like google hangouts etc., IMs in mobiles are somewhat differentfrom desktop computers. The mobile devices are limited in resources or power of conventional desktopcomputers and they therefore use memory differently. Mobiledevices do not cache data in the same way asdesktops;hence, retrieving chats are much more difficult in mobile devices. If we are recording the IMs we can get all the chats. However, it is very difficult looking at the power and other limitations.Logging the activities on client device might help but finding a complete conversationin memory is almost impossible unless chat logging is enabled.

2.6 EMAIL FORENSIC TOOLS

Various software tools have been developed to assist e-mail forensicinvestigation. These include eMailTrackerPro(http://www.emailtrackerpro.com/), EmailTracer (http://www.cyber forensics. in), Adcomplain(http://www.rdrop.com/users/billmc/adcomplain.html), Aid4Mail Forensic(http://www.aid4mail.com/), AbusePipe(http://www.datamystic.com/ abusepipe.html), AccessData's FTK (www.accessdata.com/), EnCase Forensic (http://www.guidancesoftware.com), FINALeMAIL(http://finaldata2. com), Sawmill-GroupWise (http://www.sawmill.net), Forensics Investigation Toolkit (FIT)(http://www.edecision4u. com/FIT.html), Paraben (Network) E-mail Examiner(http://www.paraben.com/email-examiner.html), etc. These analyse headers of emailmessages to detect the IP address of the originating machine. These toolsoften have abuse reporting features, e-mail classification option, support multipleencryption techniques like Credant, SafeBoot, Utimaco, EFS, PGP, GuardianEdge, Sophos Enterprise and S/MIME. Its current supported e-mail types are:Lotus Notes NSF, Outlook PST/OST, Exchange EDB, Outlook Express DBX,Eudora, EML (Microsoft Internet Mail, Earthlink, Thunderbird, Quickmail, etc.),Netscape, AOL and RFC 833. Some of these claim to be vetted by courts asstandard digital investigation platforms.
We will discuss eMailTracker Pro and EmailTracer in little detail.

2.6.1 eMailTrackerPro[34]

Email tracking is a method for monitoring the email delivery to intended recipient. Most tracking technologies use some form of digitally time-stamped record to reveal the exact time and date that an email was received or opened, as well the IP address of the recipient.

Email tracking is useful when the sender wants to know if the intended recipient actually received the email, or if they clicked the links. However, due to the nature of the technology,

[34] http://www.emailtrackerpro.com

email tracking cannot be considered an absolutely accurate indicator that a message was opened or read by the recipient.

eMailTrackerPro Standard lets you trace email back to its source, while also scanning each email message to filter out spam and harmful payloads.

Using information contained in the email header, eMailTrackerPro Standard can effectively locate the city or town that an email originated from, including Whois information that you can use to report abuse and shut them down for good. The procedure is as follows:

1. **Trace an email using the header:** To make the best use of eMailTrackerPro it's important to trace the email header, and not the email address. An email address, such as anyone@hotmail.com will just run a trace on hotmail.com, and every single time you'll get the same result. An email header is a virtual footprint telling the user where an email has travelled. Each step along the way is recorded. Spammers often try and remove/add lines to confuse where it was sent from. eMailTrackerPro can pick up on patterns and inconsistencies and mark the email as suspected spam, this isn't an exact science so anomalies can occur. An example header can be seen on the right, split up into separate lines for understanding purposes.
2. **Report Abuse:** Abuse reporting is a useful feature for users that want to take a more proactive approach to dealing with spam. EmailTrackerPro provides a platform that auto-generates an abuse report and opens a new email (may not work for all email clients) with the 'to' address filled out to the email spam address detected (as shown on the right). Once the abuse report has been sent to the email provider it is then up to them to take the next steps to shut the account down. Each account that gets shut down is one more step closer to stopping spam in the long run!
3. **Spam Filter:** The most valuable feature is the ability to trace more than one IP address or domain name at a time. Trace as many IP addresses and domain names as required and either output the results to a new tab or an Excel/HTML file.

2.6.2 Online EMailTracer

Resource Centre for Cyber Forensics (RCCF) is a pioneering institute, pursuing research activities in the area of Cyber Forensics. The centre was dedicated to the nation by the then Honorable union minister in August 2008. EmailTracer developed in RCCF is a tool to track email sender's identity. It analyzes the email header and gives the complete details of the sender like IP address, which is key point to find the culprit and the route followed by the mail, the Mail Server, details of Service Provider etc. EmailTracer traces up to Internet Service Provider level only. Further tracing can be done with the help of ISP and law enforcement agencies. The message-id will be useful for analyzing the mail logs at ISP.

2.7 SUMMARY

1. Laws nowadays give importance to emails and review them with lot of attention.
2. Email services can be Web-based email, POP3 email services, The Internet Message Access Protocol (IMAP), MAPI email servers. Most widely used protocol in emailing is simple mail transfer protocol (SMTP).

3. Few email attacks or crimes are Flaming, Email spoofing, Email bombing, Email hacking, Spams, Email frauds and Email phishing.
4. Email privacy is the broad topic dealing with issues of unauthorized access and inspection of electronic mail.
5. Emails information which will be interesting to the investigators are Email header, Body of Emails, The information hidden in the email packets and Attachments.
6. Email forensics involves capturing, securing and analysing and reporting the email evidences. E-mail forensics aims to study the source and contents of e-mail messages for evidence.
7. Various software tools have been developed to assist e-mail forensic investigation. These include eMailTrackerPro, EmailTracer.

2.8 CHECK YOUR PROGRESS

1. Fill in the blanks.

a) An Internet email message consists of three components, ____,_____,and____.
b) Many Message Transfer Agents used to accept messages for any recipient on the Internet and do their best to deliver them. Such MTAs are called_____.
c) ____ occurs when a person sends a message with angry or antagonistic content.
d) eMailTrackerPro Standard is a tool which lets you trace email back to its _____.
e) IM stands for _____.

2. State True or False

a) The Post Office Protocol 3 (POP3) provides features to manage a mailbox from multiple devices. (F)
b) Although electronic mail servers and other mail transfer agents use SMTP to send and receive mail messages, for receiving messages, client applications usually use either POP3 or IMAP. (T)
c) MUA stands for multiple user access. (F)
d) Email bombing is the intentional sending of large volumes of messages to a target address. (T)
e) Precedence in email header is used to set prioritization of queued email wheras references are used to set perspective of replies. (T)
f) Vouch by Reference (VBR) is a protocol for adding third-party certification to email. The VBR-info header contains the third party information which can be very important in tracing the email. (T)
g) Forensic analysis aims at discovering the history of a sender rather than the message. (F)

2.9 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) The message envelope, the message header, and the message body.
b) Open mail relays.
c) Flaming
d) Source

e) Instant Messaging.

2. State True or False

a) (F)
b) (T)
c) (F)
d) (T)
e) (T)
f) (T)
g) (F)

2.10 MODEL QUESTIONS

1. Describe the structure of SMTP messaging with a neat diagram.
2. Which headers in SMTP useful in tracing a message sender identity?
3. List and describe atleast 4 email attacks.
4. How is privacy a big issue in emailing?
5. What are the various types of email services?

2.11 FURTHER READINGS

1. Debra Littlejohn Shinder, Michael Cross, Scene of the Cybercrime, syngress
2. Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
3. Gutiérrez, Carlos A., Web Services Security Development and Architecture: Theoretical and Practical issues, IGI Global, 2010.

References, Article Source & Contributors

[1] Email - Wikipedia, the free encyclopedia, https://en.m.wikipedia.org/wiki/Mail_headers

[2] Email privacy - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Email_privacy

[3] Email tracking - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Email_tracking

[4] E-mail: Message Format | World4Engineers, world4engineers.com/e-mail-message-format/

[5] EMailTracer, http://www.cyberforensics.in/OnlineEmailTracer/index.aspx

[6] M. Tariq Banday, Techniques and Tools for Forensic Investigation of E-Mail, International Journal of Network Security & Its Applications (IJNSA), Vol.3, No.6, November 2011

[7] Phishing - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Phishing

UNIT III: MOBILE DEVICE FORENSICS

3.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand Mobile communication and need for mobile forensics.
- Know various ways of capturing evidences from a mobile device.
- Understand various forms of logs in mobile.
- Know the process of mobile forensics.
- Learn few mobile forensic acquisition tools.

3.2 INTRODUCTION

Mobile device forensics is a branch of digital forensics relating to recovery of digital evidence or data from a mobile device under forensicallysound conditions. The phrase mobile device usually refers to mobile phones; however, it can also relate to any digital device that has both internal memory and communication ability, including PDA devices, GPS devices and tablet computers.

The use of phones in crime was widely recognized for some years, but the forensic study of mobile devices is a relatively new field, dating from the early 2000s. A proliferation of phones (particularly smartphones) on the consumer market caused a demand for forensic examination of the devices, which could not be met by existing computer forensics techniques[1].

Mobile devices can be used to save several types of personal information such as contacts, photos, calendars and notes, SMS and MMSmessages. Smartphones may additionally contain video, email, web browsing information, location information, and social networking messages and contacts.

There is growing need for mobile forensics due to several reasons and some of the prominent reasons are:

- Use of mobile phones to store and transmit personal and corporate information
- Use of mobile phones in online transactions
- Law enforcement, criminals and mobile phone devices.

3.3 CHALLENGES IN MOBILE FORENSICS

Evidential and technical challenges exist. For example, cell site analysis following from the use of mobile phone usage coverage, is not an exact science. Consequently, whilst it is possible to determine roughly the cell site zone from which a call was made or received, it is not yet possible to say with any degree of certainty, that a mobile phone call emanated from a specific location e.g. a residential address.

To remain competitive, original equipment manufacturers frequently change mobile phone form factors, operating system file structures, data storage, services, peripherals, and even pin

connectors and cables. As a result, forensic examiners must use a different forensic process compared to computer forensics.

- Storage capacity continues to grow thanks to demand for more powerful "minicomputer" type devices.
- Not only the types of data but also the way mobile devices are used constantly evolve.
- Hibernation behaviour in which processes are suspended when the device is powered off or idle but at the same time, remaining active.

As a result of these challenges, a wide variety of tools exist to extract evidence from mobile devices; no one tool or method can acquire all the evidence from all devices. It is therefore recommended that forensic examiners, especially those wishing to qualify as expert witnesses in court, undergo extensive training in order to understand how each tool and method acquires evidence.

3.4 MOBILE COMMUNICATION

Mobile communication means can be categorized in basically three modes:

a) 802.11or WiFi
b) Bluetooth
c) Infrared(IrDA)

3.4.1 802.11 or WiFi

A wireless network is any type of computer network that uses wireless data connections for connecting network nodes.Wireless networking is a method by which homes, ecommunications networks and enterprise (business) installations avoid the costly process of introducing cables into a building, or as a connection between various equipment locations. Wireless telecommunications networks are generally implemented and administered using radio communication. This implementation takes place at the physical level (layer) of the OSI model network structure. Wi-Fi (or WiFi) is a local area wireless computer networking technology that allows electronic devices to connect to the network. Wi-Fi is the name of a popular wireless networking technology that uses radio waves to provide wireless high-speed Internet and network connections. A common misconception is that the term Wi-Fi is short for "wireless fidelity," however this is not the case. Wi-Fi is simply a trademarked phrase that means IEEE 802.11x.The Wi-Fi Alliance, the organization that owns the Wi-Fi registered trademark term specifically defines Wi-Fi as any "wireless local area network (WLAN) products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards." Initially, Wi-Fi was used in place of only the 2.4GHz 802.11b standard; however, the Wi-Fi Alliance has expanded the generic use of the Wi-Fi term to include any type of network or WLAN product based on any of the 802.11 standards, including 802.11b, 802.11a, dual-band, and so on, in an attempt to stop confusion about wireless LAN interoperability.

3.4.2 Bluetooth

Bluetooth is defined as being a short-range radio technology (or wireless technology) aimed at simplifying communications among Internet devices and between devices and the Internet. It also aims to simplify data synchronization between Internet devices and other computers.

Bluetooth products i.e. products using Bluetooth technology must be qualified and pass interoperability testing by the Bluetooth Special Interest Group prior to release. Bluetooth's founding members include Ericsson, IBM, Intel, Nokia and Toshiba.A new version of the Bluetooth wireless device-to-device technology that offers significantly lower powerconsumption than previous versions. Also referred to as Bluetooth Low Energy, Bluetooth 4.0 achieves its reduced power consumption by enabling devices to remain paired, or connected to each other, without requiring a continual stream of data to be transferred between the devices.

3.4.3 Infrared (IrDA)

Short for Infrared Data Association, a group of device manufacturers that developed a standard for transmitting data via infrared light waves. Increasingly, computers and other devices (such as printers) come with IrDA ports. This enables you to transfer data from one device to another without any cables. For example, if both your laptop computer and printer have IrDA ports, you can simply put your computer in front of the printer and output a document, without needing to connect the two with a cable.IrDA ports support roughly the same transmission rates as traditional parallel ports. The only restriction on their use is that the two devices must be within a few feet of each other and there must be a clear line of sight between them.

3.5 EVIDENCES IN A MOBILE DEVICE

As mobile device technology advances, the amount and types of data that can be found on a mobile device is constantly increasing. Evidence that can be potentially recovered from a mobile phone may come from several different sources, including handset memory, SIM card, and attached memory cards such as SD cards.

Traditionally mobile phone forensics has been associated with recovering SMS and MMS messaging, as well as call-logs, contact lists and phone IMEI/ESN information. However, newer generations of smartphones also include wider varieties of information; from web browsing, Wireless network settings, geo-location information (including geo-tagscontained within image metadata), e-mail and other forms of rich internet media, including important datasuch as social networking service posts and contactsnow retained on smartphone 'apps'.

3.5.1 Service provider logs

Although not technically part of mobile device forensics, the call detail records (and occasionally, text messages) from wireless carriers often serve as "back up" evidence obtained after the mobile phone has been seized. These are useful when the call history and/or text messages have been deleted from the phone, or when location-based services are not turned on. Call detail records and cell site (tower) dumps can show the phone owner's location, and whether they were stationary or moving (i.e., whether the phone's signal bounced off the same side of a single tower, or different sides of multiple towers along a particular path of travel). Carrier data and device data together can be used to corroborate information from other sources, for instance, video surveillance footage or eyewitness accounts; or to determine the general location where a non-geo-tagged image or video was taken.

The European Union requires its member countries to retain certain telecommunications data for use in investigations. This includes data on calls made and retrieved. The location of a mobile phone can be determined and this geographical data must also be retained. In the United States, however, no such requirement exists, and no standards govern how long carriers should retain data or even what they must retain. For example, text messages may be retained only for a week or two, while call logs may be retained anywhere from a few weeks to several months. To reduce the risk of evidence being lost, law enforcement agents must submit a preservation letter to the carrier, which they then must back up with a search warrant.

3.5.2 Subscriber identification module

A subscriber identity module (SIM) is a smart card inside of a GSM cellular phone that encrypts voice and data transmissions and stores data about the specific user so that the user can be identified and authenticated to the network supplying the phone service. The SIM also stores data such as personal phone settings specific to the user and phone numbers. If the phone not uses SIM cards then the identity information is stored in the phone hardware itself.This identification information can be used to trace a victim using service provider logs.

3.5.3 Mobile Logs

Mobile phones many a times are capable to maintain logs of calls that were made, missed and received. This information can be crucial forensically. Other logs that are also maintained mostly in the background are GPSinformation, connection information, andetc. Using these we can track the locations of mobile phones quite easily.

3.5.4 Phone books/contact lists

Phonebook names and numbersoften give investigative leads to potential witnesses and victims.Phone book can have typical information such as e-mailaddresses, home addresses, phone numbers, profile photographs, and even alternative phonenumbers.

3.5.5 Text messages

Text messages can have bits of evidence as well as date and time stamps, which can be very valuable to investigators. Often deleted messages can be recovered along with time stamps and can be used into establishing leads in an investigation.

3.5.6 Application files

Nowadays smart phones etc. have an operating system and the applications installed on these operating systems maintain lots of files and data logs which can be vita sometimes during forensic investigations.

Other forensically important data sources in a mobile devices can be Calendars and event's organizers, E-mail, Instant messages, Photos, Audio recordings etc.

3.6 MOBILE FORENSIC PROCESS

a) Seizure
b) Acquisition

c) Analysis

3.6.1 Seizure

Seizing mobile devices is covered by the same legal considerations as other digital media. Mobiles will often be recovered switched on.As the aim of seizure is to preserve evidence, the device will often be transported in the same state to avoid a shutdown, which would change files. In addition, the investigator or first responder would risk user lock activation.

However, leaving the phone on carries another risk the device can still make a network/cellular connection. This may bring in new data, overwriting evidence. To prevent a connection, mobile devices will often be transported and examined from within a *Faraday cage* (or bag). Even so, there are two disadvantages to this method. First, it renders the device unusable, as its touch screen or keypad cannot be used. Second, a device's search for a network connection will drain its battery more quickly. While devices and their batteries can often be recharged, again, the investigator risks that the phone's user lock will have activated. Therefore, network isolation is advisable either through placing the device in Airplane Mode, or cloning its SIM card (a technique which can also be useful when the device is missing its SIM card entirely).At all costs, you must keep new data from contaminating the mobile deviceafter it has been seized, for a couple of reasons.

Mobile devices can be isolated in many ways; following ways can be used to isolate a mobile on seizure:

a. **Isolating its wireless features:** By using a Faraday bag or a jamming device mobile phones can be isolated to network till the battery drains completely.Devices increase their strength to search a network; this drains the battery very fast
b. **Switch off the device:** This method is fine however, on switching on the phone lock or sim lock can be activated which can lead the phone unusable. Unlocking can be possible but is quite tricky.
c. **Airplane mode:** Airplane mode is a setting available on many mobile phones and other electronic devices that, when activated, suspends many of the device's signal transmitting functions, thereby disabling the device's capacity to place or receive calls or use text messaging – while still permitting use of other functions that do not require signal transmission (e.g., games, built-in camera, MP3 player). When the "airplane mode" is activated, it will disable all cellular services (GSM, UMTS, LTE) as well as other signal-transmitting technologies such as Wi-Fi and Bluetooth. Wi-Fi and Bluetooth can be enabled separately even while the device is in airplane mode.

3.6.2 Acquisition

The second step in the forensic process is acquisition, in this case usually referring to retrieval of material from a device (as compared to the bit-copy imaging used in computer forensics). Due to the proprietary nature of mobiles it is often not possible to acquire data with it powered down; most mobile device acquisition is performed live. With more advanced smartphones using advanced memory management, connecting it to a recharger and putting it into a faraday cage may not be good practice. The mobile device would recognize the network disconnection and therefore it would change its status information that can trigger the memory

manager to write data. Most acquisition tools for mobile devices are commercial in nature and consist of a hardware and software component, often automated.

Acquiring data from mobile phones can be very tricky and need slot of training and expertise. The acquisition can vary from mobile device to mobile device. Devices, such as cameras, are treated as storage devices in much the same way as USB drives. Mobile phones, require specific forensic software tools to extract data in a forensic way. Basic guidelines while handling digital forensic data is to be careful and see that the data on the original media is not altered in any way either by chance or intentionally. Secondly, we need to document every aspect of the investigation. Most importantly, we need to keep things centralized with proper responsibility attached to all investigators and companies involved.

Fundamentally we are looking into three components in a mobile device they are Read only Memory (ROM), Random Access Memory (RAM) and Data Storage. These components and their forensics can be very similar to that of windows or operating system forensics as discussed in Block II.

Acquiring data from mobile phones can be very tricky and need slot of training and expertise. The acquisition can vary from mobile device to mobile device. Devices, such as cameras, are treated as storage devices in much the same way as USB drives. Mobile phones, require specific forensic software tools to extract data in a forensic way. Basic guidelines while handling digital forensic data is to be careful and see that the data on the original media is not altered in any way either by chance or intentionally. Secondly, we need to document every aspect of the investigation. Most importantly, we need to keep things centralized with proper responsibility attached to all investigators and companies involved.

Fundamentally we are looking into three components in a mobile device they are Read only Memory (ROM), Random Access Memory (RAM) and Data Storage. These components and their forensics can be very similar to that of windows or operating system forensics as discussed in Unit II.

Acquisition involves following things to be done:

a) Type of Cellular Network, Code Division Multiple Access (CDMA), Global System for Mobile Communication (GSM), Integrated Digital Enhanced Network (iDEN) (A proprietary system, developed by Motorola, that uses advanced SIM cards (USIMs) and is expected to replace both CDMA and GSM).
b) Manufacturer Information of the mobile phone can be identified by Logos, Serial numbers, manufacturing codes (like IMEI: International Mobile Equipment Identifier) etc. It is advisable to cross verify the facts through Internet from online databases of the manufacturer or contact the manufacturers.
c) Phone characteristics of the device can be found from the manufacturer advertisements blogs etc. The characteristics can also guide us find areas for initial search for evidence. Some of these characteristics can be Operating system, Wireless access

methods (Bluetooth, WiFi, or infrared),Camera, manufacturer applications, internet access methods, messages etc.

3.6.3 Examination and Analysis

Mobile phone forensics analysis involves the technical examination of mobile phones and the retrieval of data from these devices. Data for analysis can be obtained from SIM cards, memory cards and from the phone handset itself. Forensic analysis of mobile phones can be carried out on various forms of data, including textual (SMS Messages),Graphic (Images), Audio Visual (Videos) and Audio (Sound recordings). Rapid advancements in mobile phone technology and the introduction of smart phones to the market by companies such as Apple and Blackberry providing large storage capacities has meant that increasingly, larger amounts of personal information is now being stored on these devices. Individuals are now becoming increasing reliant on their mobile phones as part of their daily lives. The variety of applications and facilities these devices provide including Internet, Wi-Fi, email, document viewing and editing software along with the more common mobile phone features of phonebook, call history, text messaging, voice mail, built in camera and audio facilities have seen it overlap with computer technology. The existing generation of mobile phones is sophisticated and increasingly difficult to examine however they can ultimately provide valuable evidence in prosecuting individuals. Quite often the information obtained from a phone, after intensive analysis techniques proves to be adequate for a conviction of a criminal by detectives involved with the case. Internal memory and external memory as well as the call and text records can all be analyzed to gain an insight into the activities of the mobiles owner as well as who they have been speaking or exchanging messages with. The area is ever expanding and allows for cutting edge technology to be used to keep up with the evergrowing array of mobile phones on the market today and the ever increasing feature list of these phones. Mobile forensic analysis will continue to be a specialised field while technology progresses rapidly with the sheer number of phones to be examined posing a challenge for the police.

3.7 FORENSIC ACQUISITION TOOLS

There are two categories of forensics acquisition tools. They are:

a) Hardware acquisition tools.
b) Software acquisition tools.

3.7.1 Hardware acquisition tools

We will require certain hardware to carry out acquisition. Some of the important ones are:

a. Faraday bag:A Faraday bag keeps a mobile device from communicating with an external wireless device, by intercepting radio waves and effectively acting as a large, external antenna that redirects the radio energy away from the device. Faraday bags work to keep data from reaching the mobile device and keep the mobile device from transmitting any data outward. A Faraday bag can be as small as the device you're isolating to as large as a tent when you need to do field work and need to isolate the device and your acquisition equipment at the same time. In the mobile forensic environment, isolating the device is of prime importance when you arrive on-scene.

The last thing you need is the device synchronizing on its own by way of a wireless link and changing all kinds of data.

b. **SIM card reader:**Found in any computer supply store, a card reader is used to read SIM and USIM cards without having to use the handset. Some card readers are built into the computer platform, and other card readers use a USB interface.
c. **Cable connections:**With the multitude of mobile devices now on the market, having just one mobile device connector seriously hampers your ability to do an investigation. Different mobile device manufacturers have not only different data cable connections but also different power connection interfaces. At the top of your list should reside the standard USB cable followed by the USB cable with a mini-USB connection.

3.7.2 Software acquisition tools

Certain software tools which are quite helpful while acquisition are:

a. www.MobileForensicsCentral.com: This web site provides access to a comprehensive database of phones supported by various software suppliers. A user of the web site can enter a model of a phone and the site will return a detailed report of which software and cables support it, as well what information can be retrieved from the device with the software. The goal of the site is to enable users to more efficiently find the right tool for the device they are confronted with.
b. BITPIM[35]: Allows you to view and manipulate data on LG vx4400/vx6000 and many SANYO sprint cell phones. This includes the phonebook, calendar, wallpapers, ringtones (functionality varies by phone) and the filesystem for most QUALCOMM CDMA chipset based phones.
c. CELLDEK[36]: The revolutionary celldek has been developed in cooperation with the UK's forensic science service. The portable celldekacquires data from over 200 of the most popular cell phones and PDA's. Built to perform in the field (not just in the lab), investigators can immediately gain acces to vital information, saving days of waiting for a report from a crime lab.
d. Cell Seizure[37]: Cell seizure allows you to acquire, analyze, and report on cell phone data for certain models of GsmSim Cards, Nokia, Samsung, Motorola, Sony-Ericsson, Lg, And Siemens cell phones. It can also acquire data from CDMA/TDMA phones. Designed for computer forensic examiners, cell seizure offers complete forensic examinations that can be presented in court with md5 & sha1 hash verification, write protection, html reporting, and full data dumps on some models. Version 3.0 adds support for LG, updates model support for other manufacturers, and updates sim card support.
e. Mobilyze[38]: Mobilyze is a mobile data triage tool, designed to give users immediate access to data from iOS and Android devices. Specifically designed with ease of use in mind, Mobilyze was built to respond to the mounting backlogs of evidentiary mobile devices in law enforcement agencies, both domestically and overseas. The Mobilyze

35 http://bitpim.sourceforge.net/
36 http://www.logicubeforensics.com/products/hd_duplication/celldek.asp
37 http://www.download.com/Paraben-s-Cell-Seizure/3000-2092_4-10373543.html
38 http://www.teeltech.com/

application runs on either Mac or Windows and can be effectively deployed in the field or within a forensics lab. Once Mobilyze has been installed, simply plug the smartphone or tablet into a USB port, and Mobilyze will begin collecting all relevant user data. This data is then available for viewing, searching, and filtering within minutes. Through its incredibly simple and intuitive user experience, Mobilyze allows users of all technical abilities to quickly ascertain whether a device contains relevant forensic evidence, whether immediate action needs to be taken, and/or whether the device needs to be sent to a forensics lab for a comprehensive analysis. Once relevant data is discovered, Mobilyze provides one-click reporting in a clean and easily readable format. If further analysis is required, users can seamlessly import Mobilyze data into BlackLight for a more comprehensive forensic analysis.

f. Oxygen Phone Manager II (Forensic Version)[39]: A special software for police departments, law enforcement units and all government services that wish to use the power of Oxygen Phone Manager IIfor investigation purposes. Forensic edition secures phone data to remain unchanged during extraction and exporting.
g. Oxygen Phone Manager II[40]: Oxygen phone manager ii offers management for phonebook, call register, calendar, todo lists, SMS and MMS messages, logos, tones, GPRS and WAP settings, profiles, phone dictionary, FM stations, Java games and applications.
h. Paraben's SIM Card Seizure[41]: Paraben's SIM card seizure takes the SIM card acquisition and analysis components from paraben's cell seizure and puts it into a specialized SIM card forensic acquisition and analysis tool. SIM card seizure includes the software as well as a forensic SIM card reader. If you already have cell seizure & the cell seizure toolbox, there's no need for you to get SIM card seizure as well because they contain the components to perform a forensic SIM card acquisition and analysis. This tool is for the investigator who only wants to acquire SIM cards and does not want to perform forensic exams of all cell phone data.
i. Paraben's PDASeizure[42]: Paraben's PDA seizure is a commercially available forensic software toolkit that allows forensic examiners to acquire and examine information on PDA s for both the pocket pc (PPC) and palm OS platforms 4. Paraben's product currently supports palm os up to version 5, pocket pc 2000-2003 (up to Windows CE 4.2), activesync 3.7, and hotsync. PDA seizure's features include the ability to acquire a forensic image of palm OS, pocket PC, and Blackberry devices, to perform examiner-defined searches on data contained within acquired files, generate hash values of individual files and to generate a report of the findings. PDA seizure also provides book-marking capabilities to organize information, along with a graphics library that automatically assembles found images under a single facility, based on the graphics file extension of the acquired files.

39 http://www.opm-2.com/forensic/
40 http://www.opm-2.com/OPM2/
41 http://www.paraben-forensics.com/catalog/product_info.php?cPath=25&products_id=289
42 http://www.paraben-forensics.com/handheld_forensics.htm

j. The forensicsim toolkit[43]: The forensicsim toolkit gives today's law enforcement agencies the capability to safely and confidently recover digital evidence from GSM SIM and 3G USIM devices. Acquisition, analysis and reporting form the three key stages of the forensically sound process that will save critical time and provide a cost effective solution to SIM card examinations. As an increasing number of mobile devices use high-level file systems, similar to the file systems of computers, methods and tools can be taken over from hard disk forensics or only need slight changes.

Different software tools can extract the data from the memory image. One could use specialized and automated forensic software products or generic file viewers such as any hex editor to search for characteristics of file headers. The advantage of the hex editor is the deeper insight into the memory management, but working with a hex editor means a lot of handwork and file system as well as file header knowledge. In contrast, specialized forensic software simplifies the search and extracts the data but may not find everything. Since there is no tool that extracts all possible information, it is advisable to use two or more tools for examination.

3.8 SUMMARY

1. Forensic study of mobile devices is a relatively new field.
2. There is growing need for mobile forensics due to several reasons like personal information in mobile devices; criminals as well as law agencies use mobile devices, online transactions with mobile devices.
3. Original equipment manufacturers frequently change mobile phone that's why forensic examiners must use a different forensic process compared to computer forensics.
4. Mobile communication means can be categorized in 802.11or WiFi, Bluetooth, Infrared(IrDA)
5. Mobile phone forensics involves recovering and analyzing SMS and MMS messaging, call-logs, contact lists and phone IMEI/ESN information, web browsing, Wireless network settings, geo-location information, e-mail and other forms of rich internet media such as social networking, Service provider logs, application files etc.
6. Mobile Forensic process involves seizure, acquisition, analysis. The seizure and acquisition are relatively different than that in windows system. Analysis is more like any other digital forensic analysis.
7. Forensic analysis of mobile phones can be carried out on various forms of data, including textual (SMS Messages),Graphic (Images), Audio Visual (Videos) and Audio (Sound recordings).
8. Forensic acquisition tools can be categorized in Hardware acquisition tools and Software acquisition tools. For example faraday's bag is a hardware acquisition tool whereas CellDek and CellSeizure are software tools for acquisition.

3.9 CHECK YOUR PROGRESS

1. Fill in the blanks.

[43] http://www.radio-tactics.com/forensic_sim.htm

a) _____ is a local area wireless computer networking technology that allows electronic devices to connect to the network.
b) Bluetooth 4.0 is also referred to as _____.
c) By using _____ mobile phones can be isolated to network till the battery drains completely.
d) _____ is usually referring to retrieval of material from a device for analysis and keeps its integrity intact so that they can be accepted as evidence in the court of law.
e) Various types of cellular network are: _____, _____, _____.
f) _____ work to keep data from reaching the mobile device and keep the mobile device from transmitting any data outward.

2. State True or False.

a) Bluetooth is defined as being a long-range radio technology (or wireless technology) aimed at simplifying communications among Internet devices and between devices and the Internet.
b) Network isolation is advisable either through placing the device in Airplane Mode, or cloning its SIM card.
c) Devices, such as cameras, are not treated as storage devices in much the same way as USB drives.
d) Celldek has been developed in cooperation with the UK's forensic science service.
e) Call detail records and cell site (tower) dumps cannot show the phone owner's location

3.10 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) Wi-Fi (or WiFi).
b) Bluetooth Low Energy.
c) Faraday bag or a jamming device.
d) Acquisition.
e) Code Division Multiple Access (CDMA), Global System for Mobile Communication (GSM), Integrated Digital Enhanced Network (iDEN).
f) Faraday bags.

2. State True or False.

a) False.
b) True.
c) False.
d) True.
e) False.

3.11 FURTHER READINGS

1. Andrew Hoog, Android Forensics: Investigation, Analysis, and Mobile Security for Google Android, syngress, Elsevier, 2011
2. Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.

3. I.I. Androulidakis, Mobile Phone Security and Forensics: A Practical Approach, Springer Science & Business Media, 2012.
4. Li, Chang-Tsun, Crime Prevention Technologies and Applications for Advancing Criminal Investigation, IGI Global, 2012
5. Curran, K., Robinson, A., Peacocke, S., Cassidy, S. (2010) Mobile Phone Forensic Analysis, International Journal of Digital Crime and Forensics, Vol. 2, No. 2, pp:, April-May 2010, ISSN: 1941-6210, IGI Pub.
6. Linda Volonino, Reynaldo Anzaldua; Computer Forensics for Dummies, Wiley Publishing, Inc.
7. Mobilyze, http://www.teeltech.com
8. Oxygen Phone Manager II (Forensic Version), http://www.opm-2.com/forensic/
9. Oxygen Phone Manager II, http://www.opm-2.com/OPM2/
10. Paraben's PDA Seizure, http://www.paraben-forensics.com/handheld_forensics.html
11. Paraben's SIM Card Seizure, http://www.paraben-forensics.com/catalog/product_info.php?cPath=25&products_id=289
12. The forensicsim toolkit, http://www.radio-tactics.com/forensic_sim.htm.
13. BITPIM, http://bitpim.sourceforge.net/
14. Cell Seizure, http://www.download.com/Paraben-s-Cell-Seizure/3000-2092_4-10373543.html
15. CELLDEK, http://www.logicubeforensics.com/products/hd_duplication/celldek.asp

3.12 MODEL QUESTIONS

a) What are the major sources of evidences in a mobile device? Explain.
b) Describe the mobile forensic process.
c) What are the different mobile device logs important during mobile acquisition?
d) How text messages be analysed in forensics?
e) Explain various types of mobile communications and relate this to forensic investigation.
f) What are the various ways in which mobile devices can be isolated?
g) Write the steps involved in mobile acquisition.

References, Article Source & Contributors

[1] Mobile device forensics - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Mobile_device_forensics
[2] What is Bluetooth? Webopedia, www.webopedia.com, Reproduced with permission. Copyright 1999-2015 QuinStreet, Inc. All rights reserved.
[3] What is IrDA? A Webopedia Definition, www.webopedia.com, Reproduced with permission. Copyright 1999-2015 QuinStreet, Inc. All rights reserved.
[4] What is subscriber identity module? A Webopedia Definition, www.webopedia.com, Reproduced with permission. Copyright 1999-2015 QuinStreet, Inc. All rights reserved.
[5] MOBILE/PDA FORENSIC TOOLS – Securitytools, securitytools.wikidot.com/mobile-pda-forensic-tools

UNIT IV: INVESTIGATIVE REPORTS, EXPERT WITNESS AND CYBER REGULATIONS

4.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Understand the importance of forensic reports and expert witness.
- Know basic structure and do's and don'ts while preparing a forensic report.
- Know the legal aspects in computing and cyber laws in India and abroad.
- Know the basic laws in IT Act of India.
- Know the basic offences categorized in IT Act and amendments.

4.2 INTRODUCTION

One of the most important considerations that a Forensic investigators needs to make while investigating is how to render and communicate the information gathered to the intended audience. The investigator needs to have a best approach of rendering or reporting the findings in a manner that would be categorical, technically sound and yet easily readable and understandable. A good technical report would facilitate the judicial process. A poor technical report would hamper the process and at many times induces lots of ambiguities which can lead to the acquittal of the culprit.

As in [5] Digital forensic reports can be produced for investigative purposes, separately from reports designed for litigation or electronic discovery. Oftentimes, E-Investigations reports on facts for internal review and investigation. Who used this laptop and for what purpose? Who hacked the server? Was the hacker based inside our organization or did the attack come from outside the network?

An expert witness is a very powerful source of evidence in court. Reports on data electronically discovered by computer forensics methods are important because they provide strong evidence in court documents and in overall analysis in an active lawsuit or settlement. An expert witness is one who allegedly has specialized knowledge relevant to the matter of interest, which knowledge purportedly helps to either make sense of other evidence, including other testimony, documentary evidence or physical evidence (e.g., a fingerprint). An expert witness may or may not also be a percipient witness, as in a doctor or may or may not have treated the victim of an accident or crime. In a court proceeding, a witness may be called (requested to testify) by either the prosecution or the defense. The side that calls the witness first asks questions, in what is called direct examination. The opposing side then may ask their own questions in what is called cross-examination. In some cases, redirect examination may then be used by the side that called the witness, but usually only to contradict specific testimony from the cross-examination. An expert report is a study written by one or more authorities that states findings and offers opinions.

In law, expert reports are generated by expert witnesses and investigators offering their opinions on points of controversy in a legal case, and are typically sponsored by one side or the other in litigation, in order to support that party's claims. The reports state facts, discuss details, explain reasoning, and justify the experts' conclusions and opinions.

4.3 REPORT PREPARATION

4.3.1 Gathering the data

It is highly important that the investigator has a right approach and proper planning with perspective of the case and the report that he/she is going to furnish after all findings being made. The investigator should have a priori view or idea about what form he is going to use while presenting. The right form of report would enhance the acceptability and adaptability of the audience towards right direction in the case. Documentation should be disciplined and organized. Also,this well prepared approach is essential to successful forensic technical writing. Every aspect should be well written so that it is easily understandable to all. It is advised not to use too many shortcuts or short hands as it can bring you into wrong foot many times while comprehending and can kill lot of efforts.Precise, clear and an explanative approach of writing can avoid lots of confusion to either you or the audience at a later stage. As suggested in [sans] we must be disciplined and the approach involves documenting everything elaborately as we move forward in the case investigation and findings. Also we need to keep in mind the representation of the data and findings in the final report. Thus, any need for additional forensic data will be revealed before the forensic program is completed.

4.3.2 Analyzing the results

The analysis of the result involves following steps:

a. Assumptive conclusions for lead
b. Expert report and opinion
c. Conclusive opinion
d. Data consistency and labeling
e. Hash records of the findings.

Analyzing the results is probably the most difficult because here we need to have a thorough understanding about the point and opinion that we intend to give and the point or aspects that the audience as well as the case requires. This stage overlaps the gathering data stage in the beginning parts, the initial analysis will guide us further and lead us to newer approach or ideas, thus newer set of findings and details might be required. (i.e. data analysis should begin as the data are collected). During the analysis and data review, conclusions should be drawn because the conclusions are the reason for the report and the basis for the technical report preparation. However, we must be careful while drawing conclusions. A conclusion with very few supportive data can lead to incorrectness. The incorrect conclusions are that they can create a potential for “reasonable” doubt in the courtroom. Therefore, it is best to document the conclusions in this phase (Analyzing the Results), since most of the data has already been gathered. Once the conclusions are drawn, it is advisable to list them in an order of importance with highest important conclusion first and so on.

A report offering a conclusion (an opinion) is referred to as an *expert report*. The investigator from law enforcement agencies are generally trained to merely state the facts in their reports rather giving conclusions. Once the case goes under trial a forensic analyst will be called to provide valuable suggestions. The technical witness or the forensic analyst will provide facts as found in the forensic investigation. The analyst will and can comment on the methodologies of the collection of the evidence. The forensic analyst does not offer conclusions, only the facts. However, an expert witness, (which can be another forensic analyst)can have opinions and conclusions about what was found as evidence. The opinions and conclusions are based on experience and the facts found during the forensic investigation and examination of the data obtained. Forensic analysts are usually requested to give an opinion in court about the evidences and the conclusions that can be drawn from them. In most cases, the forensic analyst's professional opinion about a case is the most useful item in any case.

It is also very important to keep data in a consistent form. i.e the records must be referenced properly with proper labels assigned to every item. Thus referring these items using labels will always help the reader to be consistent in their understanding.

Finally, we need to create MD5 hashes of the collected data/evidences and record the MD5 hashes as metadata for every file so that they can be cited in the forensic report. Creating MD5 hashes ensures the integrity of the collected data and it generates a good deal of confidence among the readers about the manner in which the investigation is being handled.

4.3.3 Outlining and organizing the report

The report outlines needs to be adhered to and the subject lines under each outline have to be addresses with proper entries and supportive write-ups. Many a times the whole case goes under confusion and jeopardy if an un-experienced forensic analyst makes a very mediocre report of the findings. A unclear and improper report can create lot of confusion leading to the loss while claiming conclusions and while performing act reconstruction. In the above two phases we needed to concentrate on what results have to be collected and reported (Gathering data and Analysis phase). In the outline phasewe need to concentrate on how the results be presented so that the conclusions can be drawn and believed into.[aa] suggests an initial template of the report and the investigator or experts can modify or adjust this according to their need. The outlines of this template are:

a. Executive Summary: Contains mainly the background of the investigation like, who authorized the forensic investigation, description of why a forensic examination of computer media was necessary, give a listing of what significant findings were found, signature block for the examiner(s) who performed theinvestigationetc. information of all people involved in the investigation along with important dates ofpertinent communications are also included.
b. Objectives: All the tasks of the investigation are outlined as well as a proper list of objectives as decided for the investigation needs to be kept here.

c. Computer Evidence Analyzed: All the evidences collected and interpreted are introduced in this section. Better way is to tabulate the evidences in form of evidence, date of collection, interpretation, expert opinion etc.
d. Relevant Findings: A summary of the findings of value are included in this section. This is the conclusions and opinions of the forensic analyst. This section tries to put the findings on the table for the reader. The reader can get an idea about what are results drawn from the evidences analyzed. It is advised to keep them in an order of increasing importance or relevance for the case.
e. Supporting Details: The findings listed in the relevant findings section is supported in a descriptive and in-depth fashion. The descriptive part suggests and emphasizes on how we reached to the conclusions in the previous section. Illustrations such as tables and figures can be very good in this section.
f. Investigative Leads: Many times because of time constraints the investigator could not proceed for further investigation though he might be having important leads. These leads can be very importantly kept in this section. The court or the client can also permit more investigation in a later stage where we can start moving further spontaneously using the leads mentioned here.
g. Subsections: In cases of computing attacks, the readers may want to know the exact attack that was performed, for this we might require analyzing a binary. So, a section "Binary Analysis" may be appropriate to the investigation. Similarly, based on requirement we can add more sub sections in other sections discussed above.

4.3.4 Writing and Revising a Rough Draft

With a logical template for computer forensic reports, writing the rough draft will be much easier. However, because many technical materials are included in forensic reports, we will be having many versions of the report. Hence we need to keep on writing a rough report and revising it. Mostly we need many reading and revising the report many times before coming to final version. Lastly we need to format the report in nice appearance using available editors. Figure 1 describes a template for a forensics report.

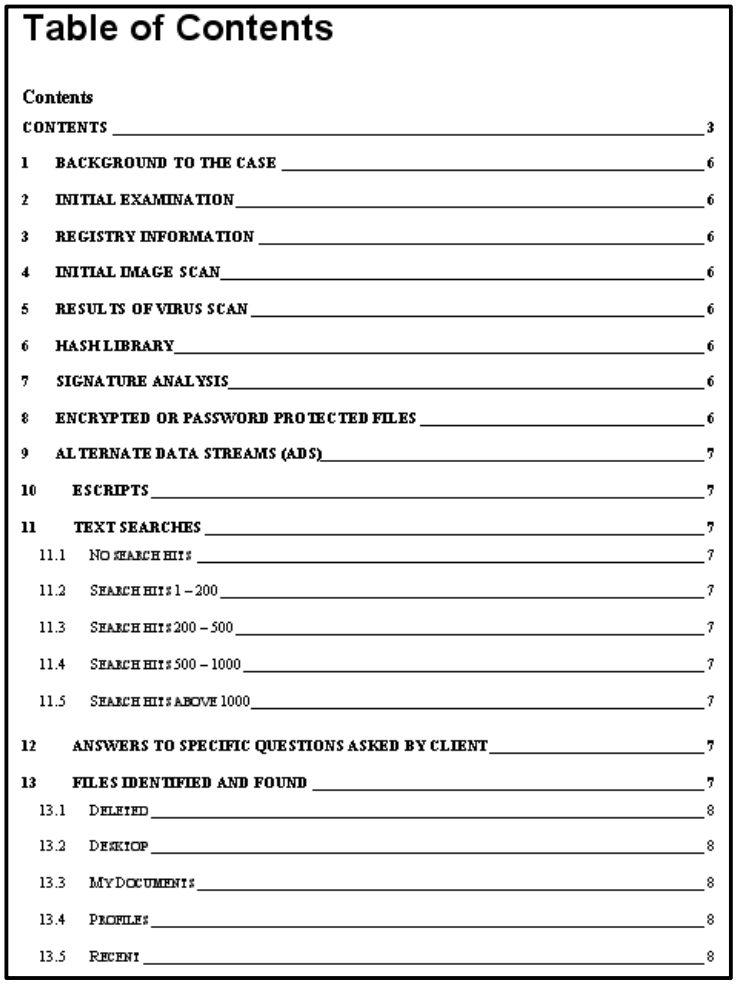

Table of Contents

Contents

CONTENTS 3

1 BACKGROUND TO THE CASE 6

2 INITIAL EXAMINATION 6

3 REGISTRY INFORMATION 6

4 INITIAL IMAGE SCAN 6

5 RESULTS OF VIRUS SCAN 6

6 HASH LIBRARY 6

7 SIGNATURE ANALYSIS 6

8 ENCRYPTED OR PASSWORD PROTECTED FILES 6

9 ALTERNATE DATA STREAMS (ADS) 7

10 ESCRIPTS 7

11 TEXT SEARCHES 7

11.1 NO SEARCH HITS 7

11.2 SEARCH HITS 1 – 200 7

11.3 SEARCH HITS 200 – 500 7

11.4 SEARCH HITS 500 – 1000 7

11.5 SEARCH HITS ABOVE 1000 7

12 ANSWERS TO SPECIFIC QUESTIONS ASKED BY CLIENT 7

13 FILES IDENTIFIED AND FOUND 7

13.1 DELETED 8

13.2 DESKTOP 8

13.3 MY DOCUMENTS 8

13.4 PROFILES 8

13.5 RECENT 8

Figure 36: A digital forensic Report format[44]

4.4 Expert Witness

In litigations, experts have become very important. Experts are involved in testimony, consultation and expert opinion.An expert witness, professional witness or judicial expert is a witness, who by virtue of education, training, skill, or experience, is believed to have expertise and specialized knowledge in a particular subject beyond that of the average person, sufficient that others may officially and legally rely upon the witness's specialized (scientific, technical or other) opinion about an evidence or fact issue within the scope of his expertise, referred to as the expert opinion, as an assistance to the fact-finder. Expert witnesses may also deliver expert evidence about facts from the domain of their expertise. At times, their testimony may be rebutted with a learned treatise, sometimes to the detriment of their reputations.

[44] Image courtesy: Computer Forensics Report Template - Privacy Resources, computer-forensics.privacyresources.org/forensic-template.htm

Typically, experts are relied on for opinions on severity of injury, degree of sanity, cause of failure in a machine or other device, loss of earnings, care costs, and the like. In an intellectual property case and an expert may be shown two music scores, book texts, or circuit boards and asked to ascertain their degree of similarity. In the majority of cases the expert's personal relation to the defendant is considered and irrelevant.

The tribunal itself, or the judge, can in some systems and call upon experts to technically evaluate a certain fact or action, in order to provide the court with a complete knowledge on the fact/action it is judging. The expertise has the legal value of an acquisition of data. The results of these experts are then compared to those by the experts of the parties.

4.4.1 Finding an expert

While finding an expert in an area of investigation we need to be very careful. Lots of people can claim to be experts in the field. It is very vital to look at the experience and expertise of an individual. The affiliation of the individual might be trivial.Apart from ability to retrieve data, a forensic analysis with expert view is more important.

The expert will likely be called to testify in court and to explain what he or she did to the computer and its data. The court will weigh the fact that theexperthad a proper training and experience, Least is the affiliation weightage in the minds of the court. The experience of an expert should be specifically in computer forensics, as skill with computers does not necessarily translate to forensic expertise. Proper consulting needs to be done with litigators who have used the expert before or have seen the expert testifying in the court.

4.4.1.1 Testifying v. Consulting

The fact that attorneys use expert witnesses for purpose of testimonies. However, nowadays in cyber cases it is becoming more common to use experts for consultations and not as testifying experts.The non-testifying experts often provide technical as well expert guidance for the attorneys to progress in the line of litigations. There can be many cases where the attorneys/investigators might not have know-how as well as skills to carry out investigation or building the case. Generally, the consulting experts need not be revealed or disclosed since a consultant is not a person having knowledge of any discoverable matter in the case.

4.4.2 What Can (and Can’t) an Expert Do?

The primary purpose of testifying experts in a given litigation is to apply scientific or technical expertise to the facts of the case and render relevant opinions that assist the trier of fact in understanding complicated or confusing matters. For instance, a forensic computer scientist will often testify to a sequence of events that took place on a given computer or network of computers. Without the expert’s testimony, the system logs, file system time stamps, and other application metadata that reveal this sequence of events, is extremely difficult to compile and present effectively. Furthermore, the expert’s special knowledge allows interpretation of the underlying data that would otherwise be inadmissible. Whereas a forensic computer expert might be able easily to determine a sequence of events that took place on a given computer, it is sometimes much harder to connect those events with a particular individual. What if the computer at issue in a case is accessible by many people?

What if the opposing party contends he was not “at the keyboard” when a pertinent event took place? A forensic computer scientist may be able to provide circumstantial evidence regarding the party who appeared to be using the computer. This might be based on the user logged in to the system. It could also be indicated by something like an individual’s Web-based email session simultaneously open at the time of other events. Since these events are less tied to the forensic computer scientist’s domain of expertise, establishment of the party using aa computer at a given time may need to be established by other means. Bringing in a computer expert for consultation early on can be extremely beneficial. For example, consider the issue of preservation. Every case an attorney is involved with carries with it a duty to preserve potentially relevant evidence. When that evidence is stored on a computer, the method of preservation becomes critical. The first issue an expert can guide you through is to explain the different preservation options available for electronically stored information (ESI). The safest option is generally forensic imaging of the storage media on the relevant computers. Forensic images are bit-for-bit copies of an entire storage medium, including space on the medium that may not currently hold any active files. This differs from simply copying all of the files on a given medium, since the inactive sections of the image may contain portions of previously deleted files, files that are still discoverable. So the first thing your expert can do is save you from falling victim to under-preservation by making an inadequate copy. After the consulting expert has explained the effectiveness of different preservation mechanisms, that expert can then further explain the impact of such preservation on your client’s computer systems. For example, large servers may be in near constant use, and require special provisions may need to be made prior to acquiring a forensic image from them. Furthermore, their storage systems may be complicated or very large, which can necessitate a greater time of unavailability. On the other hand, forensic imaging of laptop or desktop computers can often be accomplished in only a few hours, often with little or no interruption to the client’s use. These issues are difficult to navigate without a firm grasp of the underlying technology, the specialized knowledge a consulting expert can bring to bear.

4.4.3 Why Use an Expert

An effective attorney knows the facts of the case. That need to grasp the facts of the case is the key reason why an attorney should use an expert. In a trade secret case, the attorney must prove that protected information was unlawfully obtained. How can that be done if the trade secret is a customer list stored in an Excel spreadsheet? An expert can help obtain access to the computing equipment of the opposing party through discovery, and potentially find that the spreadsheet in question was copied to a USB flash drive or burned to a CD-R. Sometimes, the expert can even find that the trade secret spreadsheet was deleted, recover it and provide the time of the deletion.

4.5 LEGAL ASPECTS OF COMPUTING

IT law consists of the law (statutes, regulations, and case law) which governs the digital dissemination of both (digitalized) information and software itself (see history of free and open-source software) and legal aspects of information technology more broadly. IT law covers mainly the digital information (including information security and electronic commerce) aspects and it has been described as "paper laws" for a "paperless environment".

Cyber law or Internet law is a term that encapsulates the legal issues related to use of the Internet. It is less a distinct field of law than intellectual property or contract law, as it is a domain covering many areas of law and regulation. Some leading topics include internet access and usage, privacy, freedom of expression, and jurisdiction.

In various countries, areas of the computing and communication industries are regulated, often strictly by government bodies.

There are laws on censorship versus freedom of expression, rules on public access to government information, and individual access to information held on them by private bodies. There are laws on what data must be retained for law enforcement, and what may not be gathered or retained, for privacy reasons.

In certain circumstances and jurisdictions, computer communications may be used in evidence, and to establish contracts. New methods of tapping and surveillance made possible by computers have wildly differing rules on how they may be used by law enforcement bodies and as evidence in court.

4.5.1 Jurisdiction

Jurisdiction is an aspect of state sovereignty and it refers to judicial, legislative and administrative competence. Although jurisdiction is an aspect of sovereignty, it is not coextensive with it. The laws of a nation may have extraterritorial impact extending the jurisdiction beyond the sovereign and territorial limits of that nation. This is particularly problematic as the medium of the Internet does not explicitly recognize sovereignty and territorial limitations. There is no uniform, international jurisdictional law of universal application, and such questions are generally a matter of conflict of laws, particularly private international law. An example would be where the contents of a web site are legal in one country and illegal in another. In the absence of a uniform jurisdictional code, legal practitioners are generally left with a conflict of law issue.

Another major problem of cyberlaw lies in whether to treat the Internet as if it were physical space (and thus subject to a given jurisdiction's laws) or to act as if the Internet is a world unto itself (and therefore free of such restraints). Those who favor the latter view often feel that government should leave the Internet community to self-regulate. John Perry Barlow, for example, has addressed the governments of the world and stated, "Where there are real conflicts, where there are wrongs, we will identify them and address them by our means. We are forming our own Social Contract. This governance will arise according to the conditions of our world, not yours. Our world is different".

With the internationalism of the Internet, jurisdiction is a much more tricky area than before, and courts in different countries have taken various views on whether they have jurisdiction over items published on the Internet, or business agreements entered into over the Internet. This can cover areas from contract law, trading standards and tax, through rules on unauthorized access, data privacy and spamming to more political areas such as freedom of speech, censorship, libel or sedition.

In practical terms, a user of the Internet is subject to the laws of the state or nation within which he or she goes online. Thus, in the U.S., Jake Baker faced criminal charges for his e-conduct, and numerous users of peer-to-peer file-sharing software were subject to civil lawsuits for copyright infringement. This system runs into conflicts, however, when these suits are international in nature. Simply put, legal conduct in one nation may be decidedly illegal in another. In fact, even different standards concerning the burden of proof in a civil case can cause jurisdictional problems. For example, an American celebrity, claiming to be insulted by an online American magazine, faces a difficult task of winning a lawsuit against that magazine for libel. But if the celebrity has ties, economic or otherwise, to England, he or she can sue for libel in the British court system, where the standard of "libelous speech" is far lower.

4.5.2 Net neutrality

Network neutrality is the principle that all Internet traffic should be treated equally. According to Columbia Law School professor Tim Wu, the best way to explain network neutrality is that a public information network will end up being most useful if all content, sites, and platforms are treated equally. A more detailed proposed definition of technical and service network neutrality suggests that service network neutrality is the adherence to the paradigm that operation of a service at a certain layer is not influenced by any data other than the data interpreted at that layer, and in accordance with the protocol specification for that layer.

4.5.3 Open Internet

The idea of an open Internet is the idea that the full resources of the Internet and means to operate on it are easily accessible to all individuals and companies. This often includes ideas such as net neutrality, open standards, transparency, lack of Internet censorship, and low barriers to entry. The concept of the open Internet is sometimes expressed as an expectation of decentralized technological power, and is seen by some as closely related to open-source software.

Proponents often see net neutrality as an important component of an open Internet, where policies such as equal treatment of data and open web standards allow those on the Internet to easily communicate and conduct business without interference from a third party. A closed Internet refers to the opposite situation, in which established persons, corporations or governments favor certain uses. A closed Internet may have restricted access to necessary web standards, artificially degrade some services, or explicitly filter out content.

As of 2015, India had no laws governing net neutrality and there have been violations of net neutrality principles by some service providers. While the Telecom Regulatory Authority of India (TRAI) guidelines for the Unified Access Service license promote net neutrality, they do not enforce it. The Information Technology Act, 2000 does not prohibit companies from throttling their service in accordance with their business interests. In India, telecom operators and ISPs offering VoIP services have to pay a part of their revenues to the government.

In March 2015, the TRAI released a formal consultation paper on Regulatory Framework for Over-the-top (OTT) services, seeking comments from the public. The consultation paper was criticised for being one sided and having confusing statements. It was condemned by various politicians and internet users. By 24 April 2015, over a million emails had been sent to TRAI demanding net neutrality.

4.5.4 Indian Information Technology Act(IT Act) 2000[6]

An example of information technology law is India's Information Technology Act, 2000, which was substantially amended in 2008. The IT Act, 2000 came into force on 17 October 2000. This Act applies to whole of India, and its provisions also apply to any offense or contravention, committed even outside the territorial jurisdiction of Republic of India, by any person irrespective of his nationality. In order to attract provisions of this Act, such an offence or contravention should involve a computer, computer system, or computer network located in India. The IT Act 2000 provides an extraterritorial applicability to its provisions by virtue of section 1(2) read with section 75. This Act has 90 sections.

India's The Information Technology Act 2000 has tried to assimilate legal principles available in several such laws (relating to information technology) enacted earlier in several other countries, as also various guidelines pertaining to information technology law. The Act gives legal validity to electronic contracts, recognition of electronic signatures. This is a modern legislation which makes acts like hacking, data theft, spreading of virus, identity theft, defamation (sending offensive messages) pornography, child pornography, cyber terrorism, a criminal offence. The Act is supplemented by a number of rules which includes rules for, cyber cafes, electronic service delivery, data security, blocking of websites. It also has rules for observance of due diligence by internet intermediaries (ISP's, network service providers, cyber cafes, etc.). Any person affected by data theft, hacking, spreading of viruses can apply for compensation from Adjudicator appointed under Section 46 as well as file a criminal complaint.

The original Act contained 94 sections, divided in 19 chapters and 4 schedules. The laws apply to the whole of India. Persons of other nationalities can also be indicted under the law, if the crime involves a computer or network located in India.

The Act provides legal framework for electronic governance by giving recognition to electronic records and digital signatures. The formation of Controller of Certifying Authorities was directed by the Act, to regulation issuing of digital signatures. It also defined cybercrimes and prescribed penalties for them. It also established a Cyber Appellate Tribunal to resolve disputes arising from this new law.

Commission of cybercrime may be divided into three basic groups:

- Individual
- Organization
- Society at Large

The following are the crimes which can be committed against the following groups.

4.5.4.1 Against Individual

- Harassment via Emails
- Cyber Stalking
- Dissemination of obscene material
- Defamation
- Hacking/Cracking
- Indecent Exposure

4.5.4.2 Individual Property

- Computer Vandalism
- Transmittiming a Virus
- Network Trespassing
- Unauthorized Control over Computer System
- Hacking/Cracking

4.5.4.3 Against Organisation

- Hacking & Cracking
- Possession of unauthorised Information
- Cyber- Terrorism against Government Organisation
- Distribution of Pirated Software Etc

4.5.4.4 Against Society at Large

- Pornography
- Polluting the youth through indecent exposure
- Trafficking

The Act also amended various sections of Indian Penal Code, 1860, Indian Evidence Act, 1872, Banker's Book Evidence Act, 1891, and Reserve Bank of India Act, 1934 to make them compliant with new technologies.

4.5.5 Amendments– Indian IT Act (2008)

A major amendment was made in 2008. It introduced the Section 66A which penalized sending of "offensive messages". It also introduced the Section 69, which gave authorities the power of "interception or monitoring or decryption of any information through any computer resource". It also introduced penalties for child porn, cyber terrorism and voyeurism. It was passed on 22 December 2008 which any debate in Lok Sabha. The next day it was passed by the Rajya Sabha. It was signed by the President of 5 February 2009. The following are the list of offences and penalties.

Table 4: List of offences and the corresponding penalties

Section	Offence	Description	Penalty
65	Tampering with computer source	If a person knowingly or intentionally conceals, destroys or alters or intentionally or	Imprisonment up to three years, or/and with fine

	documents	knowingly causes another to conceal, destroy or alter any computer source code used for a computer, computer programme, computer system or computer network, when the computer source code is required to be kept or maintained by law for the time being in force.	up to ₹200,000
66	Hacking with computer system	If a person with the intent to cause or knowing that he is likely to cause wrongful loss or damage to the public or any person destroys or deletes or alters any information residing in a computer resource or diminishes its value or utility or affects it injuriously by any means, commits hack.	Imprisonment up to three years, or/and with fine up to ₹500,000
66B	Receiving stolen computer or communication device	A person receives or retains a computer resource or communication device which is known to be stolen or the person has reason to believe is stolen.	Imprisonment up to three years, or/and with fine up to ₹100,000
66C	Using password of another person	A person fradulently uses the password, digital signature or other unique identification of another person.	Imprisonment up to three years, or/and with fine up to ₹100,000
66D	Cheating using computer resource	If a person cheats someone using a computer resource or communication.	Imprisonment up to three years, or/and with fine up to ₹100,000
66E	Publishing private images of others	If a person capturee, transmits or publishes images of a person's private parts without his/her consent or knowledge.	Imprisonment up to three years, or/and with fine up to ₹200,000
66F	Acts of cyber	If a person denies access to an authorised personnel to a	Imprisonment up to life.

	terrorism	computer resource, accesses a protected system or introduces contaminant into a system, which the intention of threatening the unity, integrity, sovereignty or security of India, then he commits cyberterrorism.	
67	Publishing information which is obscene in electronic form.	If a person publishes or transmits or causes to be published in the electronic form, any material which is lascivious or appeals to the prurient interest or if its effect is such as to tend to deprave and corrupt persons who are likely, having regard to all relevant circumstances, to read, see or hear the matter contained or embodied in it.	Imprisonment up to five years, or/and with fine up to ₹1,000,000
67A	Publishing images containing sexual acts	If a person publishes or transmits images containing a sexual explicit act or conduct.	Imprisonment up to seven years, or/and with fine up to ₹1,000,000
67B	Publishing child porn or predating children online	If a person captures, publishes or transmits images of a child in a sexually explicit act or conduct. If a person induces a child into a sexual act. A child is defined as anyone under 18.	Imprisonment up to five years, or/and with fine up to ₹1,000,000 on first conviction. Imprisonment up to seven years, or/and with fine up to ₹1,000,000on second conviction.
67C	Failure to maintain records	Persons deemed as intermediatary (such as an ISP) must maintain required records for stipulated time. Failure is an offence.	Imprisonment up to three years, or/and with fine.
68	Failure/refusal to comply with orders	The Controller may, by order, direct a Certifying Authority or any employee of such Authority to take such measures or cease carrying on such activities as	Imprisonment up to three years, or/and with fine up to ₹200,000

		specified in the order if those are necessary to ensure compliance with the provisions of this Act, rules or any regulations made thereunder. Any person who fails to comply with any such order shall be guilty of an offence.	
69	Failure/refusal todecrypt data	If the Controller is satisfied that it is necessary or expedient so to do in the interest of the sovereignty or integrity of India, the security of the State, friendly relations with foreign Stales or public order or for preventing incitement to the commission of any cognizable offence, for reasons to be recorded in writing, by order, direct any agency of the Government to intercept any information transmitted through any computer resource. The subscriber or any person in charge of the computer resource shall, when called upon by any agency which has been directed, must extend all facilities and technical assistance to decrypt the information. The subscriber or any person who fails to assist the agency referred is deemed to have committed a crime.	Imprisonment up to seven years and possible fine.
70	Securing access or attempting to secure access to a protected system	The appropriate Government may, by notification in the Official Gazette, declare that any computer, computer system or computer network to be a protected system. The appropriate Government may, by order in writing, authorise the persons who are	Imprisonment up to ten years, or/and with fine.

		authorised to access protected systems. If a person who secures access or attempts to secure access to a protected system, then he is committing an offence.	
71	Misrepresentation	If anyone makes any misrepresentation to, or suppresses any material fact from, the Controller or the Certifying Authority for obtaining any license or Digital Signature Certificate.	Imprisonment up to three years, or/and with fine up to ₹100,000

4.6 SUMMARY

1. Expert reports are generated by expert witnesses and investigators offering their opinions on points of controversy in a legal case
2. The right form of report would enhance the acceptability and adaptability of the audience towards right direction in the case.
3. A report offering a conclusion (an opinion) is referred to as an expert report. During the analysis and data review, conclusions should be drawn because the conclusions are the reason for the report and the basis for the technical report preparation. However, we must be careful while drawing conclusions.
4. The outlines of a typical expert report can have sections in sequence like; Executive Summary, Objectives, Analyzed Computer Evidence, Relevant Findings, Supporting Details, Investigative Leads and other Subsections.
5. We need many reading and revising the report before coming to final version and also we need to format the report in nice appearance using available editors
6. The tribunal itself, or the judge, can in some systems and call upon experts to technically evaluate a certain fact or action, in order to provide the court with a complete knowledge on the fact/action it is judging.
7. An expert witness, professional witness or judicial expert is a witness, who by virtue of education, training, skill, or experience, is believed to have expertise and specialized knowledge in a particular subject beyond that of the average person.
8. Cyber law or Internet law is a term that encapsulates the legal issues related to use of the Internet. In various countries, areas of the computing and communication industries are regulated, often strictly by government bodies.
9. An example of information technology law is India's Information Technology Act, 2000, which was substantially amended in 2008.

4.7 CHECK YOUR PROGRESS

1. Fill in the blanks.

a) An _____ is a very powerful source of evidence in court.
b) A conclusion with very few supportive data can lead to _____.
c) The IT Act, 2000 came into force on _____.
d) _____ contains mainly the background of the investigation.
e) IT Laws has been described as "____" for a "_____".

2. State True or False

a) The IT Act 2000 has 80 sections.
b) The descriptive part of the report suggests and emphasizes on how we reached to the conclusions in the previous section.
c) Without the expert's testimony, the system logs, file system time stamps, and other application metadata that reveal this sequence of events, is difficult to compile and present effectively.
d) An expert witness, professional witness or judicial expert is a witness, who has specialized knowledge in his domain.
e) Network neutrality is the principle that all Internet traffic should have same speed.

4.8 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

a) Expert witness
b) Incorrectness
c) 17 October 2000
d) Executive Summary
e) Paper laws ,paperless environment

2. State True or False

a) False.
b) True.
c) True.
d) True.
e) False.

4.9 FURTHER READINGS

1. Becoming a Forensic Investigator - SANS Institute, https://www.sans.org/reading-room/.../forensics/forensic-investigator-1453.pdf
2. Vivek Sood, Cyber Law Simplified, Tata McGraw-Hill Education, 2008
3. PavanDuggal, Cyberlaw: the Indian perspective, Saakshar Law Publications, 2002
4. Philip J. Candilis, Robert Weinstock, Richard Martinez, Forensic Ethics and the Expert Witness, Springer Science & Business Media, 2007
5. Faust F. Rossi, Eleanor M. Fox, James T. Halverson , Expert Witnesses, American Bar Association.
6. Becoming a Forensic Investigator - SANS Institute, https://www.sans.org/reading-room/.../forensics/forensic-investigator-1453.pdf
7. Computer Forensics Report Template - Privacy Resources, computer-forensics.privacyresources.org/forensic-template.htm.
8. When to Hire a Computer Expert Witness,web.interhack.com/publications/when-to-hire.pdf

4.10 MODEL QUESTIONS

1. Describe the major amendments in the INDIAN IT Act (2008). Describe some offences and the corresponding penalties.
2. Commission of cybercrime may be divided into how many groups? Describe them.
3. What do you mean by net neutrality and open internet?
4. Why the testimonies of the experts are becoming increasingly important these days?
5. Describe the various steps of report preparation in detail.

References, Article Source & Contributors

[1] Expert witness - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Expert_witness

[2] Expert witness, http://einvestigations.com/computer-forensics/expert-witness/
[3] Information Technology Act, 2000 - Wikipedia, https://en.wikipedia.org/wiki/Information_Technology_Act,_2000
[4] Legal aspects of computing - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Legal_aspects_of_computing
[5] Net neutrality - Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Net_neutrality

ISBN: 978-93-84813-94-9

View publication stats